Humor Styles and Beyond:

The Role of Individual Differences in Humor for Psychosocial Well-Being and the Importance of Construct Validity

Thesis (cumulative thesis)

presented to the Faculty of Arts and Social Sciences

of the University of Zurich

for the degree of Doctor of Philosophy

by Sonja Heintz

Accepted in the spring semester 2017

on the recommendation of the doctoral committee:

Prof. Dr. Willibald Ruch (main supervisor)

Prof. Dr. Hugo Carretero Dios

Zurich, 2018

Abstract

The present dissertation comprises four studies that employ a multi-method approach to investigate the construct validity of the scales of the Humor Styles Questionnaire (HSQ; Martin et al., 2003): Affiliative, self-enhancing, aggressive, and self-defeating. All scales except for self-defeating were found to exhibit self-other agreement and to converge with the daily humor behaviors entailed in them. Furthermore, the relevance of the four humor styles for psychosocial well-being was investigated in self- and other-reports, daily diaries, and mixed-method investigations with cognitive interviews, which include content ratings, word analyses, and the coding of facial expressions. The affiliative and self-enhancing scales and the humor entailed in them were found to relate positively to psychosocial well-being. By contrast, the self-defeating scale was found to be a combination of general negativity (like low self-esteem) and a positive humor behavior (self-directed humor). Overall, the present dissertation suggests that the self-defeating HSQ scale seems to have limited construct validity, which necessitates changes in the construct and/or the measurement. In conclusion, it is necessary to revise the HSQ or to develop other psychometrically sound humor measurements to be able to expand our knowledge on the important topic of humor and well-being.

Zusammenfassung

Die vorliegende Doktorarbeit umfasst vier Studien, die einen multimethodalen Ansatz verwenden um die Konstruktvalidität des *Humor Styles Questionnaire* (HSQ; Martin et al., 2003) zu untersuchen: Anschlusssuchend, selbsterhöhend, aggressiv und selbstherabsetzend. Alle Skalen ausser der selbstherabsetzenden zeigten Übereinstimmung zwischen Selbst- und Fremdbericht und zeigten Konvergenz mit den in ihnen enthaltenen alltäglichen Humorverhaltensweisen. Zudem wurde die Relevanz der vier Humorstile für psychosoziales Wohlbefinden untersucht, sowohl im Selbst- und Fremdbericht, in Tagebüchern und in einem Mixed-Method-Ansatz mit kognitiven Interviews, der Inhalts- und Wortanalysen sowie die Kodierung der Mimik beinhaltet. Die anschlusssuchenden und selbsterhöhenden Skalen und der in ihnen enthaltene Humor hingen positiv mit psychosozialem Wohlbefinden zusammen. Die selbstherabsetzende Skala zeigte sich hingegen als eine Kombination aus allgemeiner Negativität (wie niedriges Selbstwertgefühl) und einem positiven Humorverhalten (selbstbezogener Humor). Insgesamt legt die vorliegende Doktorarbeit nahe, dass die selbstherabsetzende HSQ-Skala über eine eingeschränkte Konstruktvalidität zu verfügen scheint und daher Anpassungen im Konstrukt und/oder der Messung erforderlich sind. Somit ist festzustellen, dass der HSQ einer Überarbeitung bedarf oder dass andere psychometrisch fundierte Messverfahren für Humor entwickelt werden müssen, um unser Wissen zu Humor und Wohlbefinden zu erweitern.

Acknowledgments

I would like to thank everyone who supported me personally and professionally. Without your help, my studies and this dissertation would not have been possible. Although I cannot list everyone who supported me throughout the years, I would like to express my gratitude to several people who played a major role in this process.

I would like to thank Willibald Ruch, who has given me the chance to be a part of his section at the University of Zurich. His continuous knowledgeable and patient support greatly helped me to conduct my research and to find my way into academia. Also the many opportunities he offered enabled me to participate and to gain insights in many interesting projects besides my dissertation. His valuable inputs, skilled leadership, and humorous world-view enhanced the quality of the present work and made the academic work truly fulfilling.

I would like to thank my co-workers, who provided a relaxed and cheerful, but also focused and productive working atmosphere, notably Jenny Hofmann, René Proyer, Tracey Platt, Doris Baumann, Fabian Gander, Lisa Wagner, Anne Berthold, and Alexander Stahlmann. I would also like to thank colleagues and collaborators for their help with my research and for the fruitful exchange of ideas, especially Hugo Carretero Dios, who provided profound feedback and who supported and encouraged me throughout my dissertation, as well as Jessica Milner Davis, with whom I've been having insightful exchanges since the Humor Summer School in 2012.

Last but not least, I would like to express my deepest thanks and gratitude to my family and friends who always supported me unconditionally. Especially my mother Elke and my grandparents Maria and Georg have always been there for me when I needed them, and they have always encouraged and stood by me to the fullest. Without them, none of this would have been possible. (Vielen Dank für alles!)

Table of Contents

List of Tables

List of Figures

GENERAL INTRODUCTION

The present dissertation investigates the construct validity of the Humor Styles Questionnaire (HSQ; Martin, Puhlik-Doris, Larsen, Gray, & Weir, 2003) and its relevance in psychosocial well-being. The dissertation thus spans across several areas in psychology: Humor styles and humor behaviors are individual-difference phenomena relevant for personality psychology, construct validity of measures is an important aspect of psychological assessment, and psychosocial well-being is studied in clinical, social, and emotion psychology, and more recently also positive psychology. The General Introduction first covers the relevant aspects from each of these areas. Second, the current state of the literature is reviewed and the gaps and open questions are highlighted. Finally, the aims of the present dissertation are introduced.

Individual Differences in the Sense of Humor

Researchers in various disciplines of psychology have been studying individual differences in humor for almost a century, with one of the earliest pioneers including Sigmund Freud (1928). Also several influential personality researchers acknowledged the relevance of humor, such as Gordon Allport (1961), Henry Murray (Wolff, Smith, & Murray, 1934), Hans-Jürgen Eysenck (1942, 1943, 1944), and Raymond Cattell (Cattell & Luborsky, 1947; Cattell & Tollefson, 1969; Luborsky & Cattell, 1947). However, despite the assumed importance of interindividual differences in humor, the field has only seen joint and long-term research efforts since the 1980s. As McGhee and Goldstein (1983) noted: “The history of humor research prior to the 1970s can also be characterized in terms of the short-term commitment to investigating humor among those who did venture out and try their hand at designing humor studies. For reasons that remain unclear, many investigators published only one or two humor studies before abandoning the area in favor of some other research domain.” (p. V).

Personality psychologists’ interest in humor likely stems not only from its prevalence in our everyday lives (like in social interactions and the media), but also from the history of the term

"humor" (see Ruch, 1998, for details). The originally Latin term "humor" (or "umor") was first used in a medical context to denote different bodily fluids. The Greek scholar Galen of Pergamon later linked these "humors" to different temperaments, a separation that was shown to be relevant for modern personality psychology (see Stelmack & Stalikas, 1991). In the 16th century, "humor" became associated with mood (i.e., one could be either in good or bad humor). It then underwent several changes in connotation, from vice to virtue.

Nowadays, humor is either defined narrowly or broadly (as an umbrella term). The narrow definition sees humor always in a benevolent fashion as one aspect of the *comic*, besides other terms such as satire, wit, or fun (see Ruch, 1998; Ruch & Heintz, 2016a; Ruch, Heintz, Platt, Wagner, & Proyer, 2018; Schmidt-Hidding, 1963; Wirth, 2017). This view connotes humor only positively, which is consistent with the everyday usage of the term. By contrast, the most prevalent usage of 'humor' in current research is as an umbrella term that includes all comical phenomena; that is, *humor* replaces the *comic* in this broad definition. These phenomena span various domains (like humor comprehension, appreciation, and production), contents (such as satire, wit, or fun), behaviors (like laughing, telling funny stories, and making fun of oneself), emotions (like amusement, exhilaration, and *schadenfreude*), motivations (like putting others down, entertaining others, or cheering oneself up), and cognitions (like resolving a punchline). Based on this broad definition of humor, the *sense of humor* can be construed as representing an umbrella term that captures "all habitual individual differences in humor" (p. 11, Ruch, 1998). The present dissertation adopts the broad usage of the terms humor and the sense of humor. When the narrow concept of humor is concerned, the term benevolent humor is employed. One more narrow aspect of the sense of humor is the main focus of the present dissertation, namely humor styles.

Humor Styles

The term "styles" has different definitions, both in personality psychology and in humor research (see Ruch & Heintz, 2016b, for detailed treatment of the different terminologies). At the most general level, a style denotes habitual and typical behavior, in contrast to maximal behavior

(performance or ability). When this classification is adopted, humor styles refer to the 'what' and 'how' of behavior. That is, humor styles would be considered all typical behaviors associated with humor.

At a more specific level (Buss & Finn, 1987), personality traits can be classified into stylistic traits (i.e., how behavior is performed) and content traits (i.e., which behavior is shown). They subdivided the stylistic traits into instrumental style traits (how people move, speak, and gesture), affective style traits (how playful vs. serious, histrionic vs. bland, and spontaneous/expressive vs. controlled/inhibited people act), and cognitive style traits (differences in information processing). More recent theories of the components of personality traits additionally emphasize the importance of desires or motivations (like goals and wishes) besides affect, observable behavior, and cognition (see Wilt & Revelle, 2015). Although not discussed by Buss and Finn (1987), it is likely that stylistic traits can additionally be expressed in terms of one's desires. For example, people could differ in how they pursue their goals and fulfill their wishes (e.g., from direct to indirect, from immediate to delayed, or from investing very little to very much effort).

Transferring these approaches to the sense of humor, humor styles can be generally defined as individual differences in how people typically and habitually perform humor behaviors, including observable behaviors that influence one's environment (instrumental), behaviors associated with one's emotions (affective), behaviors related to one's thoughts (cognitive), and behaviors related to one's desires (motivational; see Caldwell, Cervone, & Rubin, 2008, for an extensive treatment of reasons for using humor). This represents the most comprehensive view of humor styles from the perspective of stylistic traits.

Three multidimensional approaches to humor styles were described in the literature (Craik, Lampert, & Nelson, 1996; Martin et al., 2003; Ruch, 2012; Ruch et al., 2018; Schmidt-Hidding, 1963). None of them explicitly incorporated the comprehensive definitions as put forward above, and they differ in how they define humor styles. Craik et al. (1996) derived their humor styles from a personality psychology perspective, Martin et al. (2003) derived their humor styles from a clinical

psychology perspective, and Schmidt-Hidding (1963) approached humor styles from the perspective of literature studies. These three approaches are presented in the following (see also Ruch & Heintz, 2016b, for a detailed comparison).

Craik and colleagues (1996) introduced 10 styles of everyday humorous conduct, which they derived from past psychological research and humor behaviors that they observed in everyday life. The 10 styles are sorted along five bipolar dimensions, namely socially warm vs. cold, reflective vs. boorish, competent vs. inept, earthy vs. repressed, and benign vs. mean-spirited. Their humor styles concept most strongly reflects instrumental elements (e.g., how one laughs or smiles), but it also entails affective (e.g., how humor is enjoyed), cognitive (e.g., how a punchline is understood), and motivational elements (e.g., how to put others at ease with humor). Besides stylistic traits (i.e., how behavior is shown), their approach also contains some content traits (i.e., what behavior is shown). For example, "Maintains group morale through humor" and "Makes jokes about the macabre and grotesque" contain motivation and instrumental elements, respectively. Finally, some of the aspects seem to rather reflect ability than habit (i.e., maximal instead of typical behavior; e.g., "Has the ability to tell long, complex anecdotes successfully").

Craik et al. (1996) also developed the *Humor Behavior Q-Sort Deck* (HBQD) for measuring these five bipolar styles, which entails 100 non-redundant statements that need to be Q-sorted along a normal distribution from 1 (least characteristic) to 7 (most characteristic). Q-sorting results in ipsative scores, which potentially reduce response sets and socially desirable responding, and which facilitate fine-grained differentiations among each person (see Funder, Furr, & Colvin, 2000; Ozer, 1993). At the same time, interindividual differences need to be interpreted and analyzed differently than in ratings (as every person has the same overall mean and standard deviations), and the procedure is usually more effortful and time-consuming than ratings (in which each item can be scored independent of one's scores on other items). Although this approach seems to quite comprehensively cover humor styles, it has not been frequently employed in psychological humor research. According to Thomson Reuters' Web of Science database, Craik et al.'s (1996) article has

been cited 42 times (as of March 16, 2017), and five published studies empirically assessed the styles of everyday humorous conduct (Kirsh & Kuiper, 2003; Kuiper, Grimshaw, Leite, & Kirsh, 2004; Müller, & Ruch, 2011; Ruch & Heintz, 2016b; Ruch, Proyer, Esser, & Mitrache, 2011).

One of the reasons for this lack of impact might concern the Q-sort technique, which is quite uncommon in humor research. This might also explain why most studies used ratings of the 100 HBQD statements instead of Q-sorts. In turn, changing the scoring procedure might have contributed to problems with replicating the factor structure of the HBQD. For example, the number of factors empirically recovered ranged from 3–7, and the bipolarity of and differentiation among some of the styles (e.g., reflective vs. boorish, earthy vs. repressed, and mean-spirited vs. benign) has not received univocal support (Kirsh & Kuiper, 2003; Müller, & Ruch, 2011; Ruch et al., 2011). Also the self-other agreement was found to be low for reflective vs. boorish and benign vs. mean-spirited (Ruch et al., 2011), casting doubt on the construct validity of some of the HBQD styles. Thus, although being the most comprehensive and promising approach to humor styles to date, Craik et al.'s (1996) styles of humorous conduct and the HBQD have not become the prime framework in which humor styles have been studied over the last two decades.

A fundamentally different approach to humor styles stems from literature studies, which have been recently introduced to personality psychology (Ruch, 2012; Ruch & Heintz, 2016a, 2016b; Ruch et al., 2018; Schmidt-Hidding, 1963; Wirth, 2017). In this context, "styles" does not refer to differences among people, but among works (such as texts, books, movies). Schmidt-Hidding (1963) investigated English and German texts and identified eight different comic styles: Fun, (benevolent) humor, nonsense, wit, irony, satire, sarcasm, and cynicism. This approach draws on the narrow definition of humor, in which *comic* represents the umbrella term.

Relevant for personality psychology, Schmidt-Hidding also described seven different characteristics of each of these comic styles (intention/goal, object, attitude, behavior, audience, method, and language). Specifically, the person's intention/goal underlying each comic style resembles the motivation component of humor styles (e.g., arousing sympathy in benevolent humor,

spreading good mood in fun, and improving the world in satire). The person's attitude resembles the cognitive (e.g., distant and tolerant in benevolent humor, critical in satire) and affective components (e.g., tense in wit, superior in irony, playful in nonsense, derisive in sarcasm, and destructive in cynicism). Finally, the person's behavior, method, and language when showing the comic styles represent the instrumental component (e.g., understanding in benevolent humor, callous and surprising in wit, aggressive and caricatural in satire, jolly and teasing in fun, and hostile in sarcasm). Thus, this approach seems to be promising in investigating interindividual differences in humor styles in future studies. The first empirical results already supported this notion (Ruch, 2012; Ruch & Heintz, 2016a, 2016b; Ruch et al., 2018).

Martin et al.'s Humor Styles

The third approach to humor styles stems from a clinical psychology perspective. Martin et al. (2003) presented four humor styles, which they derived from the literature on "theoretical and clinical literature on the relation between humor and well-being" (p. 51). They developed a 2 × 2 dimensional conceptualization of everyday functions of humor, which should capture "most of the elements discussed in this literature." (p. 51). Thus, this approach is not meant to be comprehensive in terms of humor styles, but it should be more or less comprehensive in terms of humor and well-being. One dimension of their conceptualization is the use or function the humor styles entail (either enhancing oneself or enhancing one's relationships). The second dimension entails the "distinction between (a) humor that is relatively benign and benevolent (i.e., tolerant and accepting of both self and others), and (b) humor that is potentially detrimental or injurious, either to the self or to one's relationships with others" (p. 52).

Combining these two dimensions results in four humor styles. They are consequently defined as "relatively benign uses of humor to enhance the self (Self-enhancing) and to enhance one's relationships with others (Affiliative), use of humor to enhance the self at the expense of others (Aggressive), and use of humor to enhance relationships at the expense of self (Self-defeating)" (p. 48, Martin et al., 2003). Comparing these definitions to the comprehensive definition

of humor styles introduced previously, Martin et al.'s (2003) humor styles cover motivational (i.e., aiming at enhancing oneself or one's relationships) and affective components (i.e., benign/benevolent or detrimental/injurious). Two observations can be made regarding the terminology employed by Martin et al. (2003). First, although derived from previous literature, the names of the four humor styles are neologisms. Second, the affiliative and self-enhancing humor styles are named according to their functions or motivational components (i.e., enhancing one's relationships and enhancing oneself, respectively), while the aggressive and self-defeating humor styles are named according to their affective components (i.e., detrimental to others and detrimental to oneself, respectively). Martin et al. (2003) do not outline how the names were chosen and why they reflect different dimensions in the 2 × 2 conceptualization.

Martin et al.'s (2003) humor styles represent the most prevalent approach to humor styles. According to Thomson Reuters' Web of Science database, Martin et al.'s (2003) article has been cited 319 times (as of March 18, 2017). In line with the conceptualization of the humor styles, most studies focused on the their relationships with psychosocial well-being (see the section below on "Humor Styles and Psychosocial Well-Being"). However, also relationships with individual-difference variables such as broad personality traits (for a meta-analysis of 15 studies, see Mendiburo-Seguel, Páez, & Martínez-Sánchez, 2015) have been investigated.

Most of these studies employed the HSQ (Martin et al., 2003), a self-report measure that assesses the four humor styles with eight items each (using a seven-point Likert scale). The HSQ was developed using the construct-based scale construction approach (Jackson, 1970), involving several stages of revisions across several adult samples (mostly psychology students). The reliabilities (test-retest and internal consistencies) and the four-factor structure of the HSQ have been supported both for the original and many translated versions (for overviews, see Martin, 2015; Ruch & Heintz, 2016b; Sirigatti, Penzo, Giannetti, & Stefanile, 2014). Additionally, several derivatives of the HSQ have been recently introduced (e.g., for children and adolescents and for the workplace; Fox, Dean, & Lyford, 2013; James & Fox, 2016; Scheel, Gerdenitsch, & Korunka,

2016). These adapted versions are not the focus of this dissertation, yet the present studies are potentially relevant for these measures as well. Although Martin et al. (2003) provided initial support for the construct validity of the HSQ, thorough validation of the measure has only started recently (see the section below on "Construct Validity of the Humor Styles Questionnaire").

Two observations can be made regarding the conceptualization of the humor styles in the existing literature, which is not always described in line with Martin et al.'s (2003) model. First, the function dimension is sometimes changed into a self-directed (self-enhancing and self-defeating) vs. other-directed (affiliative and aggressive) dimension. While this change might fit to the terminology of the humor styles, it does conflict with both the original conceptualization of the humor styles and the content of the HSQ scales. Specifically, self-enhancing does not explicitly entail humor that is directed at oneself, and affiliative does not specify any target of the humor (except in one item, which entails telling funny stories about oneself). Second, the function dimension is sometimes neglected, summarizing the four humor styles as 'adaptive' (affiliative and self-enhancing) vs. 'maladaptive' (aggressive and self-defeating). This obviously disregards the positive functions that are entailed in the latter two humor styles and thus represents an oversimplification of the concepts that is not justified by Martin et al.'s (2003) original approach to humor styles.

Overlaps Among the Different Humor Styles

A few studies empirically related the HSQ to measures of the other humor styles (i.e., the styles of humorous conduct and the comic styles). Kuiper et al. (2004) employed a rating of 32 of the 100 the HBQD statements (the Humor Behavior Deck-Revised), which they derived from scale-level principal components analyses (see Kirsh & Kuiper, 2003). It includes three scales: Skilled humor (similar to the competent and benign styles), rude/bawdy humor (similar to the boorish and mean-spirited styles), and belabored humor (similar to the inept style). In a sample of 137 psychology students, they found that the affiliative and self-enhancing scales showed large correlations with skilled humor (rs = .45 and .44, respectively, ps < .01). The aggressive scale strongly related to rude humor (r = .57, p < .01), and the self-defeating scale to belabored humor (r = .49, p < .01).

Employing a non-ipsative answer format of the HBQD (the HBQD-Rating Form) in a sample of 167 adults, Ruch et al. (2011) found that the affiliative scale was highly similar to the socially warm vs. cold style of humorous conduct ($r = .73$, $p < .001$), and it also overlapped with competent vs. inept ($r = .41$, $p < .001$). The self-enhancing scale showed strong overlaps with socially warm vs. cold ($r = .46$, $p < .001$), and the aggressive scale strongly overlapped with mean-spirited vs. benign ($r = -.40$, $p < .001$). The self-defeating scale did not show any large correlations with the HBQD-scales (all rs < |.25|).

Ruch and Heintz (2016b) administered the HBQD-Rating Form in a sample of 344 adults, separating the five bipolar styles into ten unipolar scales (to avoid confounds due to the less than perfect bipolarity in some of the styles). They found strong positive relationships of the affiliative scale with socially warm ($r = .65$, $p < .001$) and competent ($r = .49$, $p < .001$), and a negative relationship with socially cold ($r = -.54$, $p < .001$). The self-enhancing scale showed large positive correlations with socially warm ($r = .48$, $p < .001$), and the aggressive scale correlated positively with earthy and mean-spirited (rs = .57, ps < .001). Finally, the self-defeating scale correlated positively with mean-spirited ($r = .45$, $p < .001$). They also investigated how much variance the scales of the HBQD-Rating Form could explain in the HSQ scales, and vice versa. None of the scales of one measure could be completely explained by the scales of the other, but large overlaps (≥ 30% explained variance) were found for the affiliative, self-enhancing, and aggressive HSQ scales, and the socially warm, socially cold, earthy, competent, and mean-spirited HBQD-Rating Form scales. Thus, the HSQ self-defeating scale was not captured by the HBQD-Rating Form, and the reflective, boorish, repressed, inept, and benign scales were not entailed in the HSQ.

The same study also reported the overlap of the HSQ scales with one-item ratings of the Comic Styles Rating Form. The latter entailed descriptions, one for each of the eight comic styles, which participants rated according to the frequency (on a five-point scale from 1 '*never*' to 5 '*very often*'). Only one large correlation was found between the HSQ aggressive scale and sarcasm ($r = .42$, $p < .001$). The HSQ aggressive scale was also the only scale that could be explained to large

extent by the eight comic styles ($R^2 = .34$, $p < .001$). Using one-item ratings (which likely have a lower reliability) might have contributed to the lower overlap between the HSQ scales and the comic styles, making these findings preliminary.

Overall, the four humor styles overlapped with the styles of everyday humorous conduct and the comic styles, yet they were not interchangeable. Affiliative was akin to socially warm, competent, and lower socially cold styles of humorous conduct. Self-enhancing was similar to the socially warm and competent styles of humorous conduct. Aggressive was similar to earthy and mean-spirited styles of humorous conduct and the comic style sarcasm. Self-defeating showed inconsistent relationships with the other humor-related styles (either to inept/belabored or to mean-spirited). However, the comic styles and half of the ten styles of humorous conduct could not be located in the four humor styles thus far. Hence, the humor styles as conceptualized by Martin et al. (2003) were not comprehensive in terms of humor styles in general. Rather, they should comprehensively represent those humor styles that are relevant to psychosocial well-being. Psychosocial well-being is introduced next, as it is a relevant criterion of the humor styles and the second central component of the present dissertation.

Components of Psychosocial Well-Being

Psychosocial well-being entails the two broad components of psychological well-being and social well-being. Different approaches to well-being using partly different terminologies have been proposed in the literature (e.g., Diener, Suh, Lucas & Smith, 1999; Huta & Waterman, 2014; Keyes, 1998; Ryan & Deci, 2001; Ryff, 1989; Waterman, 1993). Starting from the often-cited definition originally put forward by the World Health Organization in 1948, health is defined as “a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity” (World Health Organization, 2006, p. 1). Although this definition was criticized, the relevance of the three domains of physical, mental, and social health was generally supported (e.g., Huber et al., 2011). Thus, psychological well-being (or mental well-being/health) and social well-being (or

social health) encompass both the absence of psychological or social problems and the presence of psychological and social functioning.

Psychological well-being has often been separated into hedonic and eudaimonic well-being (for overviews, see Huta & Waterman, 2014; Ryan & Deci, 2001). Eudaimonic well-being (or eudaimonia) has been advocated in psychology by Waterman (1993) and Ryff (1989; see also Ryff & Singer, 2013). It was be defined as "activity in accordance with one's daimon. This is what is considered worth having in life." (Waterman, 1990, p. 40) or "feelings of personal expressiveness" (Waterman, 1993, p. 679). The focus of eudaimonia is thus in experiencing meaning in one's life. In the hedonic view (or hedonism), well-being "consists of subjective happiness and concerns the experience of pleasure versus displeasure broadly construed to include all judgments about the good/bad elements of life." (Ryan & Deci, 2001, p. 144). The study of hedonic psychology was advocated by Kahneman, Diener and Schwarz (1999), and it is most often been studied in terms of subjective well-being (e.g., Diener et al., 1999; Pavot & Diener, 2013; Ryan & Deci, 2001). Eudaimonia and hedoism can also be construed as two different principles leading to psychological well-being, or happiness (Peterson, Park, & Seligman, 2005; Seligman, 2012)

Social well-being has received comparably less attention in the psychological literature. Based on theories in sociology and social psychology, Keyes (1998) defined social well-being as "the appraisal of one's circumstance and functioning in society" (p. 122). Additionally, some approaches to psychological well-being include components of social relationships, such as positive relations with others (Ryff, 1989), positive relationships (Seligman, 2012), and domain satisfaction in the context of subjective well-being such as work, family, and one's group (Diener et al., 1999; Pavot & Diener, 2013). Similar to these theoretical overlaps, psychological and social well-being were also found to empirically relate positively to one another: For example, Lucas and Dyrenforth's review (2006) found small to medium relationships with subjective well-being across a wide range of social relationship indicators (such as number of friends and marital status). Furthermore, recent theoretical approaches (e.g., Rusk & Waters, 2015; Seligman, 2012) and

measures (e.g., Kern, Waters, Adler, & White, 2015; Su, Tay, & Diener, 2014) incorporated eudaimonic, hedonic, and social well-being aspects. The present dissertation focuses on three concepts within psychosocial well-being, namely subjective well-being, self-esteem, and maladaptive personality, which are introduced next.

Subjective Well-Being

Subjective well-being, sometimes also subsumed under the term happiness, entails the subjective judgments of one's life as a cognitive/evaluative aspect and subjective judgments of one's emotions as an affective aspect (Diener et al., 1999; Pavot & Diener, 2013). The cognitive aspect of subjective well-being consists of judgments of global life satisfaction and the satisfaction with specific domains (like work, family, health, and the self). The affective aspect of subjective well-being consists of pleasant affects (like joy and contentment) and the lack of unpleasant affects (like sadness, depression, and anxiety). Unpleasant affects have most often been studied from a clinical perspective, while pleasant affects have received more attention within the framework of positive psychology in the last two decades. Additionally, emotion psychology deals with the experience, recognition, physiology, and expression of positive (or enjoyable) and negative emotions (e.g., Campos, Shiota, Keltner, Gonzaga, & Goetz, 2013; Ekman, 2003; Fredrickson, 2001; Izard, 1971; Lazarus, 1991; Lewis, Haviland-Jones, & Barrett, 2010; Shiota et al., 2017) and this area is thus relevant for the affective aspect of subjective well-being as well.

Empirical studies supported the separability of the different components of subjective well-being and their relevance to general well-being and health (Diener, Emmons, Larsen, & Griffin, 1985; Diener et al., 1999; Diener & Chan, 2011). Several self-report questionnaires were developed and psychometrically tested to assess the components of subjective well-being. Most prevalent among them is the Satisfaction With Life Scale (Diener et al., 1985; Pavot & Diener, 1993) to assess general life satisfaction and the Positive and Negative Affect Schedule (Watson, Clark, & Tellegen, 1988) to assess positive and negative affect. Although the most prevalent measures of subjective well-being are self-reports, especially the affective components have also been

successfully measured using behavioral indicators such as spontaneous facial expressions of positive and negative emotions (e.g., Ekman, Friesen, & Ancoli, 1980; Ekman & Rosenberg, 2005; Matsumoto, Keltner, Shiota, O'Sullivan, & Frank, 2010; Shiota et al., 2017) and emotion words that are used (Bosson, Swann, & Pennebaker, 2000; Hirsh & Peterson, 2009; Yarkoni, 2010).

Self-Esteem

Self-esteem (as a trait) is defined as "the individual's positive or negative attitude toward the self as a totality" (Rosenberg, Schooler, Schoenbach, & Rosenberg, 1995, p. 141) or "an affectively laden self-evaluation" (Leary & Baumeister, 2000, p. 1). Self-esteem can be conceptualized as one global concept (e.g., Rosenberg, 1965) or as a multidimensional concept (consisting of dimensions such as self-competence and self-liking; e.g., Tafarodi, & Milne, 2002). Rosenberg et al. (1995) argued that global self-esteem might be more relevant to psychological well-being than specific aspects of self-esteem. In line with this, empirical findings supported the relationship of global self-esteem with a wide range of outcomes relevant for well-being (for overviews, see Leary, & MacDonald, 2003; Sowislo, & Orth, 2013). Also self-esteem is relevant both for subjective well-being (e.g., Diener et al., 1985) and for eudaimonic well-being (as it is similar to self-acceptance, one of the dimensions in Ryff's theory, 1989). The most prevalent measure of self-esteem is the 10-item Rosenberg Self-Esteem Scale (Rosenberg, 1965), which assesses "feelings of self-acceptance, self-respect, and generally positive self-evaluation" (Rosenberg, Schooler, & Schoenbach, 1989, p. 1008). The reliability and validity of the scale received support in many studies (e.g., Donnellan, Ackerman, & Brecheen, 2016; Zeigler‐Hill, Besser, Myers, Southard, & Malkin, 2013).

Maladaptive Personality

In recent years, personality researchers have gained renewed interest in "dark" or subclinical personality traits, such as the Dark Triad or more recently the Dark Tetrad (Paulhus, 2014; Paulhus & Williams, 2002). Conversely, the current version of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5; American Psychiatric Association, 2013) includes one chapter on

dimensions of personality disorders, in addition to the category-based personality disorders (see Krueger et al., 2011). The DSM-5 describes 25 maladaptive personality trait facets, which are organized along five trait domains (negative affectivity, detachment, antagonism, disinhibition, and psychoticism). They should reflect extreme and maladaptive versions of the Big Five personality traits. *Negative affectivity* is characterized by experiencing a variety of negative emotions frequently and intensely and by behaving according to them. *Detachment* contains avoiding social contacts and experiencing and displaying emotions (especially positive ones) less intensely. *Antagonism* refers to treating others coldheartedly and uncaringly, by using them for one's own purpose, and by feeling superior to them. *Disinhibition* entails impulsive behavior aiming at immediate rewards and disregarding possible consequences. *Psychoticism* involves patterns of thoughts and behaviors that are unusual, eccentric, odd, or improper.

Empirical research supported that these five trait domains represent maladaptive and extreme variants of the Big Five personality factors (e.g., DeYoung, Carey, Krueger, & Ross, 2016; Gore & Widiger, 2013). Specifically, negative affectivity overlaps with neuroticism (or low emotional stability), detachment with low extraversion, antagonism with low agreeableness, disinhibition with low conscientiousness, and psychoticism with openness to experiences. Conceptually, these maladaptive personality domains are relevant for both psychological (negative affectivity, disinhibition, and psychoticism) and social well-being (detachment, disinhibition, and antagonism). Two self-report questionnaires were developed to assess either the 25 maladaptive trait facets (Personality Inventory for DSM-5; Krueger, Derringer, Markon, Watson, & Skodol, 2012) or the five maladaptive trait domains (Personality Inventory for DSM-5—Brief Form—Adult with 25 items; Krueger, Derringer, Markon, Watson, & Skodol, 2013), both in patients and healthy adults.

Humor Styles and Psychosocial Well-Being

Both the conceptualization of the humor styles and the empirical findings with the HSQ highlight the interdependence of these humor styles and psychosocial well-being. Before the empirical

findings are reviewed, the theoretically expected relationships between the humor styles and psychological and social well-being are summarized (see Table 1). To this end, three sources were extracted from Martin et al.'s (2003) paper: The definitions of the humor styles, the predictions that were postulated concerning the relationships between the humor styles (pp. 53–54) and well-being, and the contents of the 32 HSQ items.

Table 1

Hypothetical Relationships Between the Four Humor Styles and Psychosocial Well-being

Source	AF	SE	AG	SD
Psychological well-being				
Definition of the humor styles (Martin et al., 2003)	0	+	+	–
Predictions by Martin et al. (2003)	+	+	0	–
Content of the Humor Styles Questionnaire	+	+	0	–
Social well-being				
Definition of the humor styles (Martin et al., 2003)	+	0	–	+
Predictions by Martin et al. (2003)	+	0	–	–
Content of the Humor Styles Questionnaire	+	0	–	+

Notes. AF = affiliative, SE = self-enhancing, AG = aggressive, SD = self-defeating; + = positive relationship, – = negative relationship, 0 = no relationship.

As shown in Table 1, predictions across the three sources (definition, predictions, and the content of the HSQ scales) were unanimous only for self-enhancing (positive relationships with psychological well-being and no significant relationships with social well-being). The affiliative humor style was unanimously positive in terms of social well-being, and mostly positive in terms of psychological well-being (with the exception of its definition). Similarly, aggressive was unanimously negative in terms of social well-being, and mostly unrelated to psychological well-

being (again except for its definition). The self-defeating humor style was unanimously negative in terms of psychological well-being. Regarding social well-being, the definition (enhancing one's relationships) and the HSQ scale (e.g., keeping others in good spirits, making others laugh) of self-defeating were positive, while the prediction set up by Martin et al. (2003) included a negative relationship with relationship satisfaction.

The following sections present the findings on the relationships between the four humor styles (as assessed by the HSQ) and psychosocial well-being. Psychological well-being, as relevant for the present dissertation, comprises subjective well-being (i.e., life satisfaction, positive affect, and lack of negative affect), self-esteem, and the DSM-5 maladaptive personality trait domains of negative affectivity, disinhibition, and psychoticism. Social well-being is covered by the trait domains of detachment, disinhibition, and antagonism. Interpretation of the effect sizes follows Gignac and Szodorai's (2016) empirical guidelines for correlations in individual difference research. They categorized correlations as |.10| (relatively small effects), |.20| (typical or medium effects), and |.30| (relative large effects).

Humor Styles and Subjective Well-Being

Table 2 presents an overview of empirical findings on the four HSQ scales and the components of subjective well-being (life satisfaction, positive affect, and negative affect) across 14 studies. As expected, the affiliative and self-enhancing scales correlated positively with life satisfaction and positive affect, and negatively with negative affect. The effects were mostly medium-sized for affiliative, and mostly large for self-enhancing. Relationships of the aggressive scale with the components of subjective well-being were mostly small and non-significant. For the self-defeating scale, the expected relationship with lower subjective well-being was found across all components, with small to medium effects for life satisfaction and positive affect, and large effects for negative affect. Thus, the theoretically assumed relationships could be supported for the all humor styles.

Table 2

Overview of Empirical Findings on the Humor Styles Questionnaire and the Components of Subjective Well-Being

	Correlations with life satisfaction				Correlations with positive affect				Correlations with negative affect			
Study	AF	SE	AG	SD	AF	SE	AG	SD	AF	SE	AG	SD
Çalisandemir & Tagay (2015) [a]	.08	.18*	-.24*	-.12*								
Cann & Collette (2014) [a c]	.15	.33*	-.10	.01	.17	.32*	-.10	-.03	-.13	-.27*	.15	.06
Dyck & Holtzman (2013) [a]	.26*	.36*	-.09*	-.15*								
Edwards & Martin (2010) [a]	.10	.26*	-.12	-.03								
Edwards & Martin (2014) [b]	.11	.24*	.06	-.19*	.23*	.34*	.08	-.07	-.06	-.03	.27*	.36*
Jovanovic (2011)	.14*	.26*	.05	-.13								
Kazarian et al. (2009) [a]	.32*	.20	.16	.05								
Kuiper et al. (2004) [b]					.11	.28*	-.08	-.13	-.20*	-.27*	-.04	.34*
Leist & Müller (2012) [a]	.18*	.33*	.02	-.08								
Stokenberga (2008) [b]					.17*	.22*	-.04	-.11	-.09	-.21*	.11	.23*
Ruch & Heintz (2013), Sample 1 [a b]	.24*	.37*	.08	-.30*	.40*	.40*	.21*	-.21*	-.20*	-.34*	-.01	.32*
Ruch & Heintz (2013), Sample 2 [a b]	.20*	.31*	.06	-.22*	.27*	.48*	-.04	-.27*	-.19*	-.28*	-.05	.26*
Ruch & Heintz (2017) [a b]	.28*	.34*	.14	-.35*	.25*	.33*	.00	-.20*	-.20*	-.35*	-.04	.34*
Zhao et al. (2014) [a]	.19*	.25*	.00	-.01								
Mean across studies	.18	.28	-.02	-.13	.23	.34	.00	-.15	-.15	-.25	.06	.28

Notes. AF = affiliative, SE = self-enhancing, AG = aggressive, SD = self-defeating.

* $p < .05$.

[a] Employed measure for life satisfaction: Satisfaction With Life Scale (Diener et al., 1985).

[b] Employed measure for positive and negative affect: Positive And Negative Affect Schedule (Watson et al., 1988).

[c] Employed measure for positive and negative affect: Modified Differential Emotions Scale (Fredrickson et al., 2003).

Humor Styles and Self-Esteem

Table 3 presents an overview of empirical findings on the four HSQ scales and global self-esteem (across 20 studies). The affiliative and self-enhancing scales showed the expected positive relationships with global self-esteem (medium to large effects), and the self-defeating scale showed the expected negative relationships with self-esteem (large effects). The aggressive scale showed mostly either non-significant or negative correlations with self-esteem, with a small negative effect on average across the studies.

Table 3
Overview of Empirical Findings on the Humor Styles Questionnaire and Self-Esteem

	Correlations with self-esteem			
Study	AF	SE	AG	SD
Edwards & Martin (2010) [a]	.05	.27*	-.20*	-.31*
Ford et al. (2016) [a]	.29*	.30*	-.19*	-.38*
Galloway (2010) [a]	.21*	.36*	-.17*	-.38*
Hiranandani & Yue (2014), Indian sample [a]	.35*	.33*	-.11	-.22*
Hiranandani & Yue (2014), Chinese sample [a]	.20*	.28*	-.01	-.27*
Leist & Müller (2013) [a]	.31*	.37*	-.01	-.22*
Kuiper et al. (2004) [a]	.23*	.27*	-.13	-.52*
Kuiper et al. (2016) [a]	.23*	.26*	.03	-.53*
Martin et al. (2003) [a]	.21*	.28*	.03	-.36*
Martin et al. (2003) [b]	.43*	.32*	.08	-.25*
Ozyesil (2012) [a]	.18*	.24*	-.11*	-.21*
Ruch & Heintz (2013), Sample 1 [a]	.34*	.40*	.07	-.32*
Ruch & Heintz (2013), Sample 2 [a]	.25*	.45*	-.15	-.41*
Saroglou & Scariot (2002) [c]	.20	.03	-.38*	.03
Stieger et al. (2011) [a]	.23*	.29*	.04	-.20*
Vaughan et al. (2014) [a]	.12*	.25*	-.10*	-.21*
Yue et al. (2014) [a]	.28*	.38*	.00	-.12
Zeigler-Hill & Besser (2011) [a]	.27*	.32*	-.03	-.26*
Zhao et al. (2012) [a]	.29*	.32*	-.32*	-.28*
Zhao et al. (2014) [a]	.29*	.33*	-.32*	-.28*
Means across studies	.25	.30	-.10	-.29

Notes. AF = affiliative, SE = self-enhancing, AG = aggressive, SD = self-defeating.

* $p < .05$.

[a] Rosenberg Self-Esteem Scale (Rosenberg, 1965).

[b] Index of Self-Esteem (Hudson, 1982).

[c] Self-Esteem Inventory (Coopersmith, 1981).

In addition, Zeigler-Hill, Besser, and Jett (2013) investigated the relationships between the self-rated HSQ scales and peer-rated self-esteem (means across 3–5 peers). They found small overlaps for the self-enhancing (.11) and self-defeating scales (-.13) with self-esteem, and no significant overlaps of the affiliative and aggressive scales with self-esteem. Thus, the self-reported findings could be replicated partially and to a smaller extent in peer-reports.

Humor Styles and Maladaptive Personality

Only one study thus far has investigated the HSQ scales with the maladaptive personality domains. Zeigler-Hill, McCabe, and Vrabel (2016) investigated the relationships between the HSQ and the maladaptive personality domains in 594 psychology undergraduates. They found mostly medium-sized negative correlations between the affiliative and self-enhancing scales with negative affectivity, detachment, and antagonism. The affiliative scale also showed a small negative correlation with psychoticism, and the self-enhancing scale a small negative correlation with disinhibition (which were statistically significant due to the large sample size, but with overlaps of < 1% not practically significant). The self-defeating scale correlated positively with all maladaptive personality factors (mostly medium effects), and the aggressive scale correlated positively with all factors except for negative affectivity (mostly medium effects). Thus, the expected relationships with psychosocial well-being were supported for the affiliative humor style, the expected relationships with social well-being (entailed in detachment, disinhibition, and antagonism) were supported for the aggressive humor style, and the relationships with psychological well-being were supported for the self-enhancing and self-defeating humor styles. The negative correlation of the aggressive scale with psychoticism was against the theoretical expectation (see Table 1), as was the positive correlation between the self-defeating scale and the maladaptive personality factors relevant for social well-being (i.e., detachment and antagonism).

Overall, the positivity of the affiliative and self-enhancing humor styles in terms of psychosocial well-being could be empirically confirmed by the correlations of the HSQ scales with the components of subjective well-being, self-esteem, and maladaptive personality. The aggressive

humor style was neutral in terms of subjective well-being, and negative in terms of self-esteem and social well-being. Its negative relationship with self-esteem, despite small, was not in line with the theoretical predictions. Further research is needed to delineate these inconsistencies (see also Martin et al., 2003). The self-defeating humor style related negatively to psychological well-being, as expected. However, the function of this humor style in terms of social well-being was not supported: Instead of positive relationships with social well-being (as entailed in the definition and the HSQ scale), negative relationships were found. Although based on only one study that related the HSQ to the DSM-5 maladaptive personality trait domains, previous findings also suggested a relationship of the self-defeating scale with hostility and aggression (see e.g., Martin et al. 2003).

Incremental Relationships of the Humor Styles with Psychosocial Well-Being

The review of the zero-order correlations between the HSQ scales and psychosocial well-being revealed mostly medium to large relationships. Also, a meta-analysis showed that the HSQ scales were related to broad personality traits with on average medium to large effects (see Mendiburo-Seguel et al., 2015). At the same time, personality traits were related to psychosocial well-being (for overviews, see Lodi-Smith, & Roberts, 2007; Soto, 2015; Steel, Schmidt, & Shultz, 2008; Stones, Worobetz, & Brink, 2011). More specifically, the affiliative and self-enhancing scales correlated positively with extraversion, agreeableness, and emotional stability, which in turn related positively to psychological well-being. Additionally, the aggressive scale related negatively to agreeableness and conscientiousness, and the self-defeating scale related negatively to emotional stability and conscientiousness, which were in turn negatively related to psychosocial well-being. Removing the influence of broad personality traits in the relationships between the humor styles and psychosocial well-being would thus reveal their unique relationships. In other words, this informs about the degree of incremental validity the HSQ scales in explaining psychosocial well-being over and above personality, and thus their additional benefit. Of course, the degree of incremental validity can be influenced by the reliability of the personality measure: The lower the reliability of the

measure (e.g., when short scales are employed), the more variance can be potentially explained by the HSQ.

Only a few studies tested the incremental validity of the HSQ in terms of psychological well-being, and none exist for the area of maladaptive personality and measures of social well-being (for an overview, see Ruch & Heintz, 2013). Three studies tested the incremental validity of the HSQ in explaining subjective well-being and self-esteem beyond personality traits (Jovanovic, 2011; Ruch & Heintz, 2013). Jovanovic (2011) found that controlling for either extraversion or neuroticism in explaining life satisfaction yielded significant positive effects for the self-enhancing scale (which were reduced from .26 to .16), while the effect for the affiliative scale became non-significant and close to zero. In two other studies (Ruch & Heintz, 2013), controlling for the Big Five personality traits reduced the relationships of the affiliative, self-enhancing, and self-defeating scales with the components of subjective well-being and self-esteem from medium to large effects to small to medium effects. They remained significant in 4 of 18 cases for the components of subjective well-being and in 3 of 6 cases for self-esteem.

Thus, the HSQ scales showed some small to medium unique relationships, especially with self-esteem. The self-defeating scale emerged as a significant predictor of self-esteem over and above personality traits in three studies, which fits to the low self-esteem that is supposed to underlie this scale (Martin et al., 2003). Overall, incremental validity was small to medium for the self-enhancing scale in terms of subjective well-being and for the self-defeating scale in terms of self-esteem. Thus, these two HSQ scales seem to have some potential to provide additional benefits in explaining psychological well-being beyond basic personality traits.

Causal Relationships Between the Humor Styles and Psychosocial Well-Being

Several of the above-mentioned studies tried to elucidate the causal relationships between the HSQ and psychosocial well-being using mediation analyses. Mediation analyses aims at identifying variables (the mediators) that explain a relationship between two other variables. In other words, “mediators speak to how or why such effects occur” (Baron & Kenny, 1986, p. 1176).

Using the humor styles as mediators, Jovanovic (2011) found that the self-enhancing scale mediated the relationship between extraversion and neuroticism with life satisfaction, respectively. Zhao et al. (2012) found that the self-enhancing scale mediated the relationship between shyness and self-esteem. Yue, Liu, Jiang, and Hiranandani (2014) found that the affiliative and self-enhancing scales mediated the relationship between self-esteem and subjective happiness. Ford et al. (2016) found that the self-enhancing scale mediated the relationships between extraversion and happiness. Also, the affiliative and self-enhancing scales mediated the relationships of locus of control, self-esteem, and optimism with happiness.

Using the humor styles as predictors, Dyck and Holtzman (2013) found that the relationship of the affiliative, self-enhancing, and self-defeating scales with depressive symptoms and life satisfaction was mediated by perceived social support. Zhao et al. (2014) found that social support and self-esteem mediated the relationships of the affiliative and self-enhancing scales with life satisfaction. Additionally, social support mediated the relationships of these two humor styles with self-esteem. Cann and Collette (2014) found that daily-measured positive affect mediated the relationship between the self-enhancing scale (measured before the mediator) and psychological distress and life satisfaction (measured after the mediator).

Overall, these studies mostly found mediating effects for the affiliative and self-enhancing humor styles, suggesting that they might play a causal role in the relationships of personality traits or specific well-being variables with psychological well-being. Additionally, social support, self-esteem, and positive affect were found to be mediators of the relationship between the self-enhancing scale (and partly also the affiliative and self-defeating scales) and life satisfaction and negative affect. This suggests that the humor styles might influence subjective well-being indirectly via other variables (such as self-esteem).

Importantly, these studies employed cross-sectional correlational designs (except for Cann & Collette, 2014), which do not permit drawing causal conclusions. Instead, they are at best suggestive of possible causal directions in the effects between humor styles, self-esteem, and

psychological well-being, which would need to be corroborated with appropriate methodological designs, such as experimental manipulations or specific longitudinal designs (Fiedler, Schott, & Meiser, 2011; Kline, 2015; Tate, 2015).

Ruch and Heintz (2013, 2017) used another approach to determine the causality underlying the relationship of the HSQ scales with psychosocial well-being. They manipulated the HSQ items to arrive at item versions that only contain the humor content (by removing the non-humor parts in each item) or they replaced the humor content with non-humorous contents to arrive at item versions that only contained the non-humor parts. Correlating these manipulated HSQ versions with psychological well-being variables informs about the relevance of the humor contents vs. non-humor parts in the HSQ for psychosocial well-being.

In the first study (Ruch & Heintz, 2013), they found that the HSQ version that only contained humor correlated positively with subjective well-being and self-esteem (for affiliative and self-enhancing) or was neutral (aggressive and self-defeating). The same pattern was found for the HSQ version that contained only the non-humor parts, with the exception that self-defeating scale was negative in terms of psychological well-being. Additionally, controlling for the non-humor parts in the relationship between the HSQ scales and psychological well-being removed the positive effect of affiliative and the negative effect of self-defeating, while self-enhancing still showed medium positive effects with positive affect and self-esteem. These findings with the non-humor version of the HSQ were recently replicated (Ruch & Heintz, 2017). This second study did not find any incremental effects of the HSQ scales beyond their non-humor parts. Two conclusions can be drawn from these studies: First, the negative relationship of the self-defeating humor style was only recovered in the non-humor parts of the scale, but not in the humor entailed in it. Second, the relationships of the HSQ with psychological well-being seemed to be mostly driven by the non-humor parts entailed in its items (like going overboard, not being concerned, enjoying, being with friends and family, effectively coping with problems), and less so by the humor content entailed in them.

One recent study used an experimental procedure to elicit humor styles to assess their effects on state anxiety (Ford, Lappi, O'Connor, & Banos, 2017). A role-play exercise was conducted in which participants anticipated taking a math test. Three between-participants conditions were employed: The self-enhancing and self-defeating humor conditions involved humorous materials of each of these humor styles and instructions to behave according to the humor styles. Compared to participants in a no-humor control condition and in the self-defeating condition, participants in the self-enhancing condition had a lower state anxiety, while no significant differences were found between the other two conditions. Ford et al. (2017) also tested whether presenting the self-enhancing humor materials alone would have an effect. Indeed, they found that participants in the self-enhancing humor materials, in comparison to the no-humor control group, had higher scores in humorous perspective taking (i.e., reframing the math test in a humorous way), while they did not have lower anxiety scores. The causal roles of affiliative and aggressive humor styles were not tested, and employing the self-defeating humor style was not found to lead to more anxiety than the other conditions. The latter was thus not found to negatively influence psychological well-being, in contrast to the correlational findings with the self-defeating scale. It is important to note that Ford et al. (2017) did not employ a pre-post design, which precludes interpretations of changes in state anxiety due to the experimental manipulations.

In general, the shared variance between psychosocial well-being and the HSQ scales (up to 11.6% on average for self-enhancing) represent mostly medium to large effects. The findings in terms of incremental validity and the experimental study by Ford et al. (2017) instead suggest that the overlap between the HSQ scales and psychosocial wellbeing might be small or negligible. The latter findings are also in line with other experimental studies on humor that found mostly small to medium effects of humor interventions on psychological well-being (e.g., Crawford & Caltabiano, 2011; Ruch, Hofmann, Rusch, & Stolz, 2018; Wellenzohn, Proyer, & Ruch, 2016a, 2016b). Thus, the correlational and self-reported findings of the relationships between the HSQ and psychological well-being might be overestimations of the 'true' covariance between humor and well-being.

Validity as a Central Psychometric Property

When constructing psychological tests, it should be ensured that they are objective, reliable, and valid. Objectivity refers to the standardization in employing the test and in analyzing and interpreting the test scores. This can, for example, be achieved by written instructions, by a response key that unambiguously links the individual answers to the scale scores, and by a description of how the scale scores should be interpreted. Reliability refers to the accuracy of measurement, or the ratio of true-score variance to observed variance. If a psychological test only contains measurement error, then reliability would be zero, and reliability would be perfect if a measure was error-free. Reliability can be computed as the internal consistency (i.e., the size of the relationships among the scale items), as test-retest-reliability (by correlating the scale scores across two points of measurement), and as parallel-test reliability (by correlating the test scores of two tests that should measure the same construct). Finally, validity refers to the extent with which the test scores represent the construct that is intended to be measured. Objectivity and reliability are necessary, but not sufficient preconditions for validity. The process of validation encompasses testing different components of validity, and construct validity can be conceptualized as an umbrella term that combines several of these components (see e.g., Cronbach & Meehl, 1955; Loevinger, 1957; Messick, 1995).

Construct Validity

In 1954, the American Psychological Association issued the *Technical recommendations for psychological tests and diagnostic techniques*, in which they distinguished among four kinds of validity studies: Concurrent validity (cross-sectional correlation of the test score with an external criterion), predictive validity (correlation of the test score with an external criterion that is measured at a later point in time), content validity (the comprehensiveness of the test items for the construct of interest, which is usually established by expert consensus), and construct validity (which constructs the test scores represent). Construct validity is always needed when constructs are to be assessed that cannot be directly observed and instead need to be inferred from observable

indicators. In this case, it is necessary to establish that the scores of the observable indicators (e.g., test items) are actually caused by the construct of interest, and not by other variables.

In their seminal paper on construct validity, Cronbach and Meehl (1955) extended the view entailed in the technical recommendations. They outlined several methods of testing construct validity, such as group differences, correlations between tests that should measure the same construct (which was later termed "convergent validity"), factor structure, experimental interventions, and investigating the response process. They also noted that construct validation is a complex process, which aims at establishing a nomological network of a test. This involves "propositions relating test to construct, construct to other constructs, and finally relating some of these constructs to observables." (p. 294). When construct validity is not supported empirically, then either the test does not measure the construct, or the conceptualization of the construct is incorrect, or the empirical study was faulty. If the latter is not the case, Cronbach and Meehl (1955) outlined the consequences for a published test as "it should be reported as a stop sign to discourage use of the test pending a reconciliation of test and construct, or final abandonment of the test" and "The test may serve, at best, only as a source of suggestions about individuals to be confirmed by other evidence" (p. 296). This emphasizes the relevance of construct validity, as it indicates what a test score actually means and consequently whether a test should be used or not.

More recent approaches to construct validity used a more encompassing definition, which combines concurrent, predictive, and content validity into construct validity as an umbrella term (see e.g., Clark & Watson, 1995; Messick, 1995; Simms, 2008). Messick (1995) also stressed the importance of the consequences derived from test scores, in addition to the meaning attached to them: "The primary measurement concern with respect to adverse consequences is that any negative impact on individuals or groups should not derive from any source of test invalidity, such as construct underrepresentation or construct-irrelevant variance" (p. 746). This cautions against misleading conclusions that could be drawn from tests lacking construct validity. For example, interventions or therapies might be derived from the empirical findings with a measure, which could

in fact be detrimental or adverse rather than beneficial if the test scores do not actually capture the construct to be measured.

Importantly, construct validation of a test is best already investigated in each stage of the test construction process. For example, Loevinger (1957) suggested testing content validity (substantial validity) during item generation, factorial validity and reliability (structural validity) during scale construction, and finally criterion, convergent, and discriminant validity (external validity) of the final scale (see also Clark & Watson, 1995; Simms, 2008; Ziegler, 2014). On a more general note, assessing constructs with multiple methods is important to avoid common method variance, which can bias the relationship among measures (see Podsakoff, MacKenzie, Lee, & Podsakoff, 2003). In the following, two prominent approaches to assessing construct validity are introduced, the multitrait-multimethod approach (based on Campbell & Fiske, 1959) and cognitive survey testing techniques (see Beatty & Willis, 2007; Meitinger & Behr, 2016; Presser et al., 2004).

The Multitrait-Multimethod Approach

Introduced by Campbell and Fiske (1959), the multitrait-multimethod (MTMM) approach represents a convenient way to comprehensively study the convergent and discriminant validity as two central aspects of construct validity. Convergent validity refers to an overlap between measures that should assess the same construct, and discriminant validity refers to a lack of overlap between measures that should assess dissimilar constructs. This extends the nomological network of the measure and suggests necessary revisions for either the measure or the construct if results do not support the expected relationships. Campbell and Fiske (1959) argued that assessing at least two constructs (two “traits” in the MTMM terminology) with at least two methods (based on different sources of information, such as self-, peer- and observer-reports) yields information on the degree of variance due to trait, method, and error entailed in the test scores. For example, the four HSQ scales could be measured by self- and peer-reports (as two different methods). The observed correlations between the self- and peer-reports of one HSQ scale represent its convergent validity (i.e., the extent to which the different methods measure the same trait), while the correlations of one

HSQ scale with another HSQ scale (both self- and peer-reported) represent their degree of discriminant validity. Although Campbell and Fiske (1959) did not present any stringent cut-off criteria, the convergent correlation of a scale should be higher than its discriminant correlations to support its construct validity.

The MTMM approach has been widely used to test construct validity, and recent statistical advances enabled a more sophisticated test of the degree of construct validity. Specifically, the MTMM approach based on the observed correlations (with at least a 2 traits × 2 methods correlation matrix) does not allow to separate method variance from error variance. Structural equation modeling allows to model latent factors instead of observed variables, enabling the separation of variance due to trait, method, and error (and depending on the model also further sources of method- and error-specific variance). Different models have been developed for different kinds of MTMM data, such as interchangeable (e.g., different peer raters) and structurally different methods (e.g., self- and peer-reports), for multilevel data, and for longitudinal data (see e.g., Eid et al., 2008; Koch, Schultze, Burrus, Roberts, & Eid, 2015; Koch, Schultze, Eid, & Geiser, 2014; Ziegler, Booth, & Bensch, 2013).

Cognitive Interviewing

Cognitive interviewing is prevalently used in the early stages of test construction, mainly in survey research and applications. It is defined as "the administration of draft survey questions while collecting additional verbal information about the survey responses, which is used to evaluate the quality of the response or to help determine whether the question is generating the information that its author intends" (Beatty & Willis, 2007, p. 287). Thus, cognitive interviews can be employed to determine the degree of construct validity of a test: The cognitive process underlying the answer to the items of a measure should reveal whether participants indeed understand the questions in line with the construct to be measured, whether their interpretations vary, or whether they systematically deviate from the intended construct. If variations or deviations occur in the interpretations, items can be adapted or discarded to arrive at a set of items that comprehensively and adequately

represents the construct. In psychology, cognitive interviewing techniques are still infrequently employed, despite their potential to improve construct validity in test construction and to investigate the construct validity of existing questionnaires. Different cognitive interviewing techniques can also be fruitfully combined with one another and with other test construction strategies (like factor analyses; see Latcheva, 2011). This represents a mixed-models approach (or method triangulation), which is similar to the MTMM approach, yet combines quantitative with qualitative methods (for an overview, see Small, 2011).

Several techniques for conducting cognitive interviews have been proposed: Thinking aloud, cognitive probing, and online cognitive probing (see Beatty & Willis, 2007; Behr, Bandilla, Kaczmirek, & Braun, 2014; Behr, Kaczmirek, Bandilla, & Braun, 2012; Presser et al., 2004). Thinking aloud requires participants to continuously report their thoughts while completing the test. This method closely captures participants' interpretations (i.e., the retrieval process underlying the answers), yet it might interfere with how participants respond to the items. For example, participants might elaborate on the items and their answer to a greater extent than they would have done in a standard test situation. Cognitive probing entails asking questions after the participants completed each item or at the end of the survey. These questions can be tailored to the test constructor's need, for example to detect comprehension problems (e.g., "How do you understand this question?"), to determine how specific terms are interpreted (e.g., "What does X mean to you?"), or to determine the process underlying an answer (e.g., "Why did you chose this answer option?"). This technique also reduces the interference of the cognitive interviewing with the response process. However, the reports are generated retrospectively and thus reveal less information about the retrieval process underlying the answers. Recently, cognitive probing was successfully applied to online testing by having participants write the answer to the probes instead of orally reporting them (Behr et al., 2012, 2014). Although this technique usually elicits shorter responses, a large sample size can be generated, thus increasing the representativeness of responses (as discussed by Meitinger, & Behr, 2016).

The present dissertation thus employs both online cognitive probing and think-aloud techniques to counterbalance the strengths and weaknesses of each technique. Combining them with quantitative analyses in a mixed-method approach represents a promising and rather novel approach to testing construct validity. Although best done during test construction, this approach can also be used to determine the degree of construct validity of existing tests, such as the HSQ.

Construct Validity of the Humor Styles Questionnaire

In comparison to the large number of studies that investigated the relationships of the HSQ with other constructs, the literature on the construct validity of the HSQ is rather limited. In the following, studies that investigated the self-other agreement of the HSQ scales (i.e., self-reports and reports of knowledgeable informants, like peers or romantic partners) and validity studies based on self-reports are presented. The criterion validity (and additionally the incremental validity) of the HSQ scales in terms of psychosocial well-being has already been reviewed previously (see the section on "Humor Styles and Psychosocial Well-Being"), as was the factorial validity of the four scales (see the section on "Martin et al.'s Humor Styles"). What has not been investigated yet is the content validity of the HSQ (as one aspect of construct validity as an umbrella term). Specifically, experts would need to evaluate to what extent the four HSQ scales comprehensively and representatively cover the everyday functions of humor described the literature on humor and well-being (for details and applications of content validation in humor research, see Carretero Dios, Pérez, & Buela Casal, 2009; Delgado Rico, Carretero Dios, & Ruch, 2012). Content validity is not covered by the present dissertation and would thus be a task for future research in the area.

Construct Validity in Terms of Self-Other Agreement

Thus far, four studies investigated the construct validity of the HSQ in terms of the agreement between self- and other-reports of the HSQ scales. First, Martin et al.'s (2003) construction article investigated the agreement of undergraduate students' self-reported HSQ scales and reports by their dating partners on one item of each HSQ scale ($N = 165$). Convergent validities were medium to

large (.22 for affiliative, .33 for self-enhancing, .32 for aggressive, and .25 for self-defeating), and discriminant validities were close to zero with the exception of affiliative (up to .18 with self-enhancing). Second, using a sample of 80 Australian students, Findlay and Jones (2005) investigated the agreement between self-, partner-, and friend-reports of the HSQ scales. They found "moderate agreement between the three judgments" (p. 204), while they reported no information on the degree of discriminant validity. Third, Cann, Zapata, and Davis (2011) investigated the agreement between self-reports and partner-reports in sample of 82 young adult couples. Overlaps were medium to large for each humor style (.28 for affiliative, .20 for self-enhancing, .24 for aggressive, and .30 for self-defeating), and they did not report the degree of discriminant validity. Fourth, Zeigler-Hill et al. (2013) investigated the self-peer agreement of the HSQ in a sample of 257 psychology undergraduates. The peer-reports of a short version of the HSQ (consisting of three items for each humor style) were aggregated across three to five peers. Employing multilevel-modeling, they found medium to large convergent validities for each humor style (.33 for affiliative, .20 for self-enhancing, .34 for aggressive, and .34 for self-defeating). Discriminant validities were low between affiliative and self-enhancing (overlaps of .21 and .17), while the other styles were not significantly related with one another across the self- and peer-reports.

Overall, these findings suggest small to large self-other agreements (i.e., convergent validity) for the affiliative and self-enhancing scales, and medium to large agreements for the aggressive and self-defeating scales. The size of these convergent correlations can be compared with the values from meta-analyses of self-other agreements of the Big Five personality traits (average correlation = .36; Connolly, Kavanagh, & Viswesvaran, 2007) and subjective well-being (average correlation = .42; Schneider & Schimmack, 2009). Additionally, self-other agreement ranged from ranged from .20 to .58 (M = .38) for the styles of humorous conduct (Ruch et al., 2011) and from .40 to .56 (M = .49) for the comic styles (Ruch et al., 2018). This shows that the convergent validities of the HSQ scales were lower than the average of similar traits.

However, the size of the convergent validities (if observed correlations are used) is limited by the reliabilities of the measures. Using Spearman's (1904) formula for correction for attenuation yields values that are independent of the size of the reliabilities. The average convergent validity corrected for attenuation was .55 for the Big Five personality traits (Connolly et al., 2007; Schneider & Schimmack, 2009 did not report corrected coefficients for subjective well-being). Additionally, Carretero Dios, Eid, and Ruch (2011) investigated the self-peer agreement of the trait version of the State-Trait-Cheerfulness-Inventory (Ruch, Köhler, & Van Thriel, 1996), which assess three dimensions of the humorous temperament, using a combination of multi-level and structural equation modeling. The shared variance for the three scales cheerfulness, seriousness, and bad mood ranged from .27 to .32, corresponding to true-score correlations of .51–.57. Regarding the HSQ, the corrected values can be computed for Cann et al.'s (2011) study only, as the other studies lack the necessary information to conduct the computations (i.e., reliabilities and zero-order correlations). The correction increased the convergent validities of the HSQ scales from on average .26 to .32 (.34 for affiliative, .24 for self-enhancing, .32 for aggressive, and .37 for self-defeating). Thus, both the uncorrected and corrected convergent validities of the HSQ scales were lower than the comparable values for the Big Five personality traits, subjective well-being, the other humor-related styles, and the temperamental basis of the sense of humor.

Discriminant validity was supported for each scale except for affiliative and self-enhancing, for which the discriminant relationships were almost as the convergent ones. This indicates that the two humor styles that are proposed to be beneficial are confounded, and more research is needed to determine how they differ from one another (especially in behaviors readily observable for others).

Taken together, the comparably low self-other agreement of the HSQ might either reflect a lack of construct validity of the HSQ scales, a lack of validity of the humor style constructs, or it might be due to problems with the studies that examined construct validity. The latter explanation can be explored by considering the design of the four studies. First, the studies by Martin et al. (2003) and Cann et al. (2011) compared the self-reports to reports of the target's partner. This

considers only one type of knowledgeable informant, and it does not take advantage of arriving at more reliable and representative other-reports by combining the judgments of several sources. Second, Martin et al. (2003) and Zeigler-Hill et al. (2013) relied on short versions for assessing the other-reports (one and three items of the HSQ, respectively), which potentially limits both the reliability and the validity of the obtained other-reports. Third, the samples employed by Findlay and Jones (2005) and Cann et al. (2011) were rather small and much smaller than the recommended sample size of at least 150 to achieve stable correlation coefficients (see Schönbrodt, & Perugini, 2013). Fourth, none of the studies employed MTMM analyses based on structural equation modeling, in which it is possible to separate trait and method variance from error variance. Implementing research designs that employ multiple other-reports, large sample sizes, the full HSQ scale, and structural equation modeling of the MTMM data are necessary to overcome the shortcomings of the existing studies. This would help to decide whether indeed the construct validity of the HSQ is compromised and revisions of the scale and/or the constructs are needed.

Construct Validity Based on Self-Reports

Two recent studies investigated the construct validity of the HSQ using self-reports. Heintz and Ruch (2015; see also Martin, 2015; Heintz & Ruch, 2016) aimed at comparing the HSQ scales to the definitions and the construct descriptions of the humor styles as entailed in the construction article by Martin et al (2003). The study aimed to test whether the humor styles are comparable across these three stages of test construction (with the definitions representing the second stage and the HSQ the final stage) as perceived by the participants. In other words, would those who score high in a HSQ scale also agree to the definitions of this scale and to the construct descriptions of this style? They employed a single-indicator correlated-traits correlated-methods minus one model (see Eid, 2000), a MTMM model based on structural equation modeling. The model entailed four trait factors (i.e., the four humor styles) and two methods (definitions and construct descriptions). The self-reported HSQ was chosen as a reference method in the model against which the other two methods were compared. In a sample of 340 adults, they found large agreements between the three

sources for all HSQ scales (≥ .47), and a small overlap was found between the self-enhancing scale and its definition (.16). The largest overlap (i.e., lowest discriminant validity) was found between affiliative and self-enhancing (.68). Thus, convergent validity was supported for all HSQ scales except for self-enhancing, and discriminant validity was supported for all HSQ scales except for affiliative and self-enhancing. This supports the notion that the conceptualization (especially the definition) or the measurement of the self-enhancing humor style might require further scrutiny.

Ruch and Heintz (2017) employed another approach by experimentally separating the construct-relevant content (i.e., humor) and construct-irrelevant non-humor parts within the HSQ items (as discussed in section "Causal Relationships Between the Humor Styles and Psychosocial Well-Being"). Correlating these manipulated versions with the original HSQ scales yields information on their convergent validity. They found that affiliative and self-enhancing were more strongly determined by their construct-relevant content than by the non-humor parts, while the reverse pattern was found for self-defeating, and the findings were mixed for aggressive. Thus, convergent validity was supported for the affiliative and self-enhancing scales, to a lower extent for the aggressive scale, and not at all for the self-defeating scale. This implies that interpreting the self-defeating scale in terms of its humor-relevant content (i.e., the humor use) might be misleading, as the scale was mainly determined by the non-humor parts entailed in its items, like going overboard, doing things more than one should, and doing things too much. This also fits to the finding that the non-humor parts, but not the humor, entailed in the scale was responsible for the negative relationships with psychological well-being (Ruch & Heintz, 2013, 2017). Overall, this highlights the need for further investigations on what the humor entailed in this scale actually reflects. Also, it shows that the effects of the humor content and non-humor parts entailed in the scale should be separated when relevant criteria such as psychosocial well-being or humor are concerned.

Aims of the Present Dissertation

The present dissertation includes three empirical research papers (Parts I–III) that test four overarching aims. The aims fill the gaps that exist in the literature and thus focus on the relationships of the HSQ scales with psychosocial well-being and the construct validity of the HSQ.

Aim 1: Testing the Degree of Self-Other Agreement of the HSQ Scales with a Sound Methodology (Part I)

As outlined in the section "Construct Validity in Terms of Self-Other Agreement", the existing studies of the self-other agreement of the HSQ scales suffered from various shortcomings, which might have contributed to the inconsistent and comparably small convergent validities found. Thus, Part I employs a large sample (i.e., $N > 150$), at least two other-reports (by knowledgeable informants), and a modern MTMM approach based on structural equation modeling. This allows testing the convergent and discriminant validity of the HSQ scales more appropriately. If self-other agreement is supported, the previous findings could be attributed to problems with the study design rather than the lack of construct validity of the HSQ scales. If convergent and/or discriminant validity is not supported, it would be necessary to call into question either the affected HSQ scale and/or the humor style construct. This would then require a reinterpretation of existing findings that employed the scale and caution against further usage of the scale in its current form (until it is revised).

Aim 2: Assessing the Relationships of the HSQ Scales and Psychosocial Well-being Beyond Self-Reports (Parts I and III)

As reviewed in the section on "Humor Styles and Psychosocial Well-Being", medium to large positive relationships were established of psychosocial well-being with affiliative and self-enhancing, and medium to large negative relationships were established with self-defeating. Also aggressive was found to show negative relationships with social well-being. As these findings were largely based on self-reports, the present dissertation extends these investigations to other-reported

HSQ scales in Part I, and to content ratings, word frequency analyses, and the frequency of facial displays of positive and negative emotions in Part III. This shows (a) whether significant relationships can be established beyond the common method variance among the measures (as is the case if both humor styles and psychosocial well-being are self-reported; see Podsakoff et al., 2003) and (b) whether the social reality of the positivity and negativity entailed in the humor styles can be supported. That is, the positive and negative associations with psychosocial well-being should not only be entailed in how the persons see themselves, but also in other people's judgments and in the spontaneous behavior they show (i.e., whether they show facial displays or use words of positive or negative emotions). This yields information on the construct validity of the HSQ scales in terms of criterion validity (relationships with psychosocial well-being). Part I focuses on the maladaptive personality domains as described in the DSM-5, and Part III focuses on self-esteem (or positive vs. negative self-evaluation) and positive vs. negative emotions as the affective components of subjective well-being as well as the improvement of relationships.

Aim 3: Delineating the Humor Entailed in the HSQ Scales (Parts II and III)

This aim extends the previous studies that separated humor from the non-humor parts in the HSQ scales (Ruch & Heintz, 2013, 2017) by extracting the humor entailed in the HSQ scales. Part II employs a daily-diary method to investigate all humor behaviors entailed in the HSQ, complemented by non-redundant humor behaviors entailed in the other two approaches to humor styles (i.e., the styles of humorous conduct and the comic styles). Investigating the overlap between the HSQ scales and their aggregated humor behaviors (across three to five days) informs about their criterion validity, their social reality, and their convergent and discriminant validity. The HSQ scales should be able to predict their aggregated daily humor behaviors, and they should do so to a larger extent than they predict the behaviors of any of the other HSQ scales.

Part III focuses on the self-defeating scale, employing cognitive interviewing techniques (online cognitive probing and thinking aloud) in a mixed-method approach. This informs about the cognitive processes underlying the answers to the self-defeating items and allows analyzing the

extent to which participants interpret self-directed humor when responding to the self-defeating items. If the self-defeating scale measured a humor style, it would need to converge with the ratings of self-directed humor.

Aim 4: Delineating the Relationships of the Humor Entailed in the HSQ Scales and Psychosocial Well-Being (Parts II and III)

Extracting the humor in the HSQ also allows testing its overlap with psychosocial well-being. Previous studies suggested that the humor entailed in the HSQ scales might not be negative or detrimental, but rather positive or neutral in terms of psychological well-being (Ford et al., 2017; Ruch & Heintz, 2013, 2017). Aim 4 extends these studies and tests the interplay of the humor entailed in the HSQ with psychosocial well-being more thoroughly. Part II relates the aggregated daily humor behaviors to subjective well-being. Additionally, the incremental relationships of the humor behaviors beyond the Big Five personality traits and the HSQ are tested to see whether any variance in subjective well-being can be uniquely explained by the humor behaviors. Part III employs content ratings of self-directed humor to test whether the amount of self-directed humor entailed in the explanations provided in the cognitive interviewing is related (a) to positive vs. negative emotions (measured via word frequency analyses and the frequency of facial displays of positive and negative emotions), (b) to content ratings of positive vs. negative self-evaluation (i.e., self-esteem), and (c) to improving one's relationship with others. These analyses show how the humor entailed in the self-defeating scale relates to affect, self-evaluation, and relationships across independent measurement methods.

PART I: Do Others Judge my Humor Style as I do? Self-Other Agreement and Construct Validity of the Humor Styles Questionnaire

Sonja Heintz

This part is an adapted version of the following article:

Heintz, S. (2017). Do others judge my humor style as I do? Self-other agreement and construct validity of the Humor Styles Questionnaire. *European Journal of Psychological Assessment*. Advance online publication. doi:10.1027/1015-5759/a000440

Abstract

Humor research has intensified in psychology over the last two decades, with the Humor Styles Questionnaire (HSQ) being the most prevalent measure. Still, the construct validity of its four scales (affiliative, self-enhancing, aggressive, and self-defeating) has not received univocal support. The present study uses a multitrait-multimethod approach to test the self-other agreement of the four HSQ scales with 202 targets and two knowledgeable informants per target. Employing a multilevel multiple-indicator correlated trait-correlated (method-1) (ML-CT-C[M-1]) model informed on the construct validity of the HSQ. Discriminant validities were sufficient for all scales. Convergent validity was supported for three of the four HSQ scales, except for the self-defeating scale. Similarly, the overlaps of the self- and other-reported HSQ scales with maladaptive personality as external criteria converged for all HSQ scales except for the self-defeating scale. Taken together, the present findings suggest that the self-defeating scale does not measure the maladaptive humor style it is supposed to measure.

Keywords: Humor Styles Questionnaire; self-other agreement; construct validity; multitrait-multimethod analysis; ML-CT-C(M-1) model

Introduction

Humor is studied in many areas of psychology. Investigating self-other agreement as one aspect of construct validity is especially important for humor measures, as humor is largely a social phenomenon. Thus, others should judge our humor similarly as we do, and deviations between the two perceptions could potentially have undesirable consequences (like failing at achieving social support or offending others with one's humor). The present study focuses on the Humor Styles Questionnaire (HSQ; Martin, Puhlik-Doris, Larsen, Gray, & Weir, 2003), as it is the most frequently used instrument to assess individual differences in the sense of humor and as validity evidence for the HSQ is scarce and partially conflicting. Investigating the self-other agreement should yield an in-depth picture of the construct validity and the social reality of the HSQ, revealing whether the scales actually measure what they are supposed to measure (i.e., different humor styles).

Multitrait-Multimethod Approach

Two aspects of construct validity are relevant for the present study (see Campbell & Fiske, 1959): Convergent validity, the extent to which self-other agreement of the HSQ scales is established, and discriminant validity, the extent to which the HSQ scales can be distinguished from one another. Both aspects of construct validity can be simultaneously tested in a multitrait-multimethod approach, which separates the variance due to trait, method, and measurement error. Recent multitrait-multimethod approaches based on structural equation models overcome the limitations of investigating observable correlations between the traits and methods by disentangling these different sources of variance (for overviews, see Eid, Lischetzke, Nussbeck, & Trierweiler, 2003; Eid et al., 2008).

Specifically, the multilevel multiple-indicator correlated trait-correlated (method-1) (ML-CT-C[M-1]) model was proposed as a fruitful option for multitrait-multimethod data involving structurally different methods (like self- and other-reports) and interchangeable

methods (like raters; Carretero Dios, Eid, & Ruch, 2011). The multilevel portion of the model contains at level 1 the raters, who are supposed to be drawn from a larger pool of possible raters of the target (and are thus interchangeable). As several raters judge the same target, they are nested within the target. Targets, including their self-reports and the average other-report, are modeled at level 2. The CT-C(M-1) portion of the model contains one latent factor for each trait in the confirmatory factor analysis. One method factor is dropped from the model, as one reference method is chosen (the self-reports in the present study) against which the other non-reference methods are compared (the other-reports in the present study). At level-2, the trait factors thus represent the variance common to the self-reported HSQ scales, and the method factors represent the deviations of the average other-report from the trait factor (i.e., the residual).

The ML-CT-C(M-1) model provides several advantages (see Carretero Dios et al., 2011; Eid et al., 2003, 2008). First, measurement error can be separated from method and trait variance, as originally envisioned by Campbell and Fiske (1959); in other words, the true-score variances of traits and methods are investigated. Second, modeling the multilevel structure of raters nested in targets allows separating potential biases of individual raters at level-1 (unique method specificity) from the potential biases of the average other-report at level 2 (common method specificity). Third, trait-specific method effects can be estimated, as method biases might often vary across the different traits. Previous applications empirically supported the usefulness of the ML-CT-C(M-1) model (e.g., Carretero Dios et al., 2011). The present study applies the model for the first time to self- and other-reports of the HSQ.

Construct Validity of the HSQ

The HSQ measures four trait-like humor styles, defined as “the interpersonal and intrapsychic functions that humor is made to serve by individuals in their everyday lives, and particularly those functions that are considered most relevant to psychosocial well-being” (Martin et al.,

2003, p. 51). Two humor styles are supposed to be adaptive (affiliative and self-enhancing), and two are supposed to be potentially maladaptive (aggressive and self-defeating).

Four studies thus far compared self and other-reports of the HSQ scales. Martin et al. (2003) investigated in the construction article the agreement of students' self-reported HSQ scales and reports by their dating partners on one item of each HSQ scale ($N = 165$). Convergent validities were small to medium for the four HSQ scales, and discriminant validities were close to zero except for a positive correlation between the affiliative and self-enhancing scales. Findlay and Jones (2005) investigated the agreement between self-, partner-, and friend-reports of the HSQ scales in 80 students and found "moderate agreement between the three judgments" (p. 204). Cann, Zapata, and Davis (2011) investigated the agreement between self-reports and partner-reports in sample of 82 couples, and they found small to medium convergent validities for each HSQ scale. Finally, Zeigler-Hill, Besser, and Jett (2013) investigated the self-peer agreement of the HSQ in 257 students. The peer-reports of three items for each HSQ scale were aggregated across several peers, resulting in small (self-enhancing) to medium convergent validities (affiliative, aggressive, and self-defeating). Again, small to medium positive relationships emerged between the affiliative and self-enhancing scales. Overall, these studies yielded small to medium convergent validities of the HSQ scales, and lower discriminant validities among the affiliative and self-enhancing scales. However, this might have been due to suboptimal methodologies used; that is, no structural equation modeling approaches were employed, and the other-reports were either measured with short scales or they were not averaged across raters.

Additionally, two recent studies investigated the construct validity of the HSQ scales using self-reports. Comparing the HSQ scales to the definitions and the construct descriptions of the humor styles by employing a single-indicator CT-C(M-1) model, large agreements were found between the three sources for all HSQ scales, and a small overlap was found between the HSQ self-enhancing scale and its definition (Heintz & Ruch, 2015). The largest

overlap (i.e., lowest discriminant validity) was found between affiliative and self-enhancing. Another approach (Ruch & Heintz, 2017) involved experimentally separating the construct-relevant content (i.e., humor) and construct-irrelevant context within the HSQ items. Correlating these manipulated versions with the original HSQ scales showed that the affiliative and self-enhancing scales were mainly determined by their construct-relevant content, while the self-defeating scale was mainly determined by the construct-irrelevant context entailed in its items (e.g., going overboard).

Overall, these studies yielded partially conflicting and preliminary results, making further investigations of the construct validity and especially the self-other agreement of the HSQ scales necessary. As the HSQ is usually studied in the context of psychosocial well-being, the present study additionally investigates whether the previous self-report findings can be replicated with the other-reports of the HSQ in terms of maladaptive personality (see Zeigler-Hill, McCabe, & Vrabel, 2016). This would further support the social reality and relevance of the HSQ for our psychosocial well-being, in addition to its construct validity.

Materials and Methods

Participants

Targets. Overall, 468 participants agreed to take part in the study, of which 306 had complete and usable scores (65.4%). Only participants who for whom two other-reports were available were considered in the final sample, which was the case for 202 (72.3% female, 27.7% male) participants. They were on average 26.37 years old (*SD* = 11.00, range 18–75 years) and they were primarily Swiss (84.2%), German (9.9%), or had another nationality (5.9%). Two-thirds of them were college or university students (59.9%), 17.3% had passed tertiary education, 16.3% had a high school diploma, and 6.4% completed an apprenticeship.

Raters. Overall, 489 raters (knowledgeable informants) agreed to take part in the study, of whom 404 (82.6%) provided complete and usable scores (56.2% female, 43.8%

male). Overall, there were 202 dyads (two independent raters for each target). They were on average 33.87 years old (*SD* = 14.87, range 18–72 years). Most raters indicated that they were a friend (38.9%) or a relative (child, sibling, or parent; 38.9%) of the target, 17.1% were romantic partners, and 5.2% indicated other types of relationships (e.g., work colleague). The raters were very familiar with the targets: The average relationship length was 14.04 years (*SD* = 9.89, range 1–60), and raters on average indicated that they knew the person very well (*M* = 6.32, *SD* = 0.82, range 3–7) on a Likert-type scale from *very little knowledge* (1) to *excellent knowledge about the person* (7).[1]

Instruments

Humor Styles Questionnaire (HSQ; Martin et al., 2003; German version by Ruch & Heintz, 2016b). The HSQ measures four humor styles with eight items each. Sample items are "I usually don't laugh or joke around much with other people" (affiliative), "If I am feeling depressed, I can usually cheer myself up with humor" (self-enhancing), "If someone makes a mistake, I will often tease them about it" (aggressive), and "I don't often say funny things to put myself down" (self-defeating). The instrument employs a seven-point Likert scale from *totally disagree* (1) to *totally agree* (7). Internal consistencies (McDonald's omega) were sufficient (from .75 for aggressive to .88 for self-enhancing).

Humor Styles Questionnaire – Other-Report Form (adapted for this study). The instrument consists of the same 32 items as the HSQ, yet they were rephrased to refer to another person instead of oneself. Specifically, the possessive pronoun "my" was replaced by

[1] The rank correlations between the two measures of familiarity (relationship length and knowledge) and accuracy (computed as the squared Euclidian distances between the self- and other-reports, separate for each HSQ scale) were small and mostly nonsignificant ($-.13 \leq \rho \leq .15$). Thus, differences in familiarity across the different target-rater dyads did not influence the accuracy of the other-reports.

"her/his" (adapted to the target's gender), and the pronoun "I" was replaced by the targets' first name. Sample items are "[Name] usually doesn't laugh or joke around much with other people" (affiliative), "If [Name] is feeling depressed, he/she can usually cheer himself/herself up with humor" (self-enhancing), "If someone makes a mistake, [Name] will often tease them about it" (aggressive), and "[Name] doesn't often say funny things to put himself/herself down" (self-defeating). It employs the same seven-point Likert scale as the HSQ. McDonald's omega of the other-reports (aggregated across two raters) was good (from .82 for aggressive to .89 for affiliative).

Personality Inventory for DSM-5—Brief Form—Adult (PID-5-BF; Krueger, Derringer, Markon, Watson, & Skodol, 2013; German version by Zimmermann, Krueger, Markon, & Leising, 2012). The PID-5-BF assesses five maladaptive personality factors (negative affectivity, detachment, antagonism, disinhibition, and psychoticism) described in the Diagnostic and Statistical Manual of Mental Disorders (American Psychiatric Association, 2013) with five items each. Sample items are "I worry about almost everything" (negative affectivity), "I often feel like nothing I do really matters" (detachment), "I use people to get what I want" (antagonism), "People would describe me as reckless" (disinhibition), "I have seen things that weren't really there" (psychoticism). Items were answered on a four-point scale ranging from *very false or often false* (0) to *very true or often true* (3). McDonald's omega was sufficient (ranging from .67 for antagonism to .73 for psychoticism).

Procedure

The study was conducted online (www.unipark.info) and in accord with the local ethical guidelines. All participants (targets and raters) declared their online informed consent. After completing the self-reports, targets were provided with a link to the HSQ Other-Report Form, which they should forward to at least two people who knew them well. The link included a

unique identifier number to match the other-reports anonymously with the self-reports. Other variables were assessed that are not relevant for the present study.

Statistical Analyses

The ML-CT-C(M-1) model was computed with MPlus 5.1 (Muthén & Muthén, 1998–2008). The model was run separately for each pairwise combination of HSQ scales to reduce the complexity of the model to achieve an optimal ratio of free parameters to the (level-2) sample size.[2] Four two-item parcels were created for each HSQ scale using a balancing approach. A power of .82 with an alpha level of .05 was achieved for two-tailed correlations of .20.

Results

Observed Self-Other and Inter-Rater Agreement

The four HSQ scales exhibited on average large inter-rater agreements (ICC_{mean} = .51, range .44–.62) and self-other agreements (r_{mean} = .48, range .31–.59). (Table 5 in the Appendix additionally shows the means and standard deviations and the agreement for each HSQ item and scale.) Additionally, item-profile agreement was computed by averaging the correlation between each self-other and each rater dyad across the eight HSQ items of each scale. Inter-rater item-profile agreements were large for affiliative (r_{mean} = .76) and self-defeating (r_{mean} = .47), medium to large for aggressive (r_{mean} = .38), and medium for self-enhancing (r_{mean} = .30). Self-other item-profile agreements were medium for self-enhancing (r_{mean} = .31 and self-defeating (r_{mean} = .30), and medium to large for affiliative (r_{mean} = .41) and aggressive (r_{mean} = .35).

ML-CT-C(M-1) Model

Table 4 shows the fit indices of the six estimated ML-CT-C(M-1) models (one for each pairwise combination of HSQ scales).

[2] Results were highly similar when the four traits were included in one model simultaneously.

Table 4

Fit Indices of the Six Multilevel Multiple Indicator Correlated Trait-Correlated (Method-1) Models

Model	χ^2	χ^2/df	CFI	TLI	RMSEA	SRMR 1	SRMR 2
AF-SE[a]	295.05	2.06	.94	.94	.05	.03	.11
AF-AG[a]	217.58	1.52	.96	.96	.04	.03	.12
AF-SD[a]	277.66	1.94	.94	.94	.05	.05	.10
SE-AG[b]	253.33	1.77	.94	.94	.04	.03	.12
SE-SD	295.65	2.07	.93	.93	.05	.04	.11
AG-SD[b]	228.89	1.60	.95	.95	.04	.03	.11

Note. df = 143. SRMR 1 = SRMR at level 1, SRMR 2 = SRMR at level 2.

[a] The models in which affiliative was included were used to determine convergent validity; results only differed slightly in the other models.

[b] The latent variable covariance matrix was not positive definite for these models.

As shown in Table 4, most models showed an overall good to accepting fit. The SRMR for level-1 was always good, while the SRMR of level-2 was unsatisfactory. Figure 1 illustrates the ML-CT-C(M-1) model including the standardized factor loadings (The means, unstandardized factor loadings, residual variances, and reliabilities of the model indicators are shown in Table 6 in the Appendix.)

Focusing on the results at level-2 (N = 202), Figure 1 shows that the loadings between the observed indicators of the self-reports ("Self 1"–"Self 4") and the latent trait factors were large for the four HSQ scales (ranging from .59 to .85). Squaring these loadings yields the level-2 reliabilities of the self-reports; they ranged from .35 for aggressive to .66 for self-enhancing (*Mdn* = .58). The latent correlations between the average other-reports (across raters), modeled as a latent variable at level-2, and the latent trait factors were high for three

of the HSQ scales (.77–.90) and medium to large for self-defeating (.46). Squaring these latent correlations yields the convergent validities at level 2, which are interpreted as the amount of variance in the average other-reports that can be explained by the self-reports. These convergent validities were high for affiliative (for which 65.6% of the variance of the average other-reports was explained by the self-reports), self-enhancing (59.3%), and aggressive (81.0%), and lower for self-defeating (21.2%).

The common method factor loadings between the average other-reports and the common method factors ("Other CM") represent the extent to which the average other-report cannot be explained by the self-reports (i.e., the residuals). These loadings ranged from .42–.89. Squaring these common method factor loadings represents the amount of variance specific to the average other-report (i.e., not shared with self-reports). They were smaller than the convergent validities at level-2, ranging from 17.6% (aggressive) to 41.0% (self-enhancing), with the exception of self-defeating (79.2%).

The observed other-reports ("Other 1"–"Other 4") at level-1 ($N = 404$) can be subdivided into four different variance components: The variance due to the unique method factor (level-1), the common method factor (level-2), the trait factor (level-2) and measurement error (level-1). The loadings between the observed other-reports and the unique method factors ("Other UM") capture the deviation of one rater from the average other-report of the same target. Squaring these loadings yields the level-1 reliabilities, which were in general lower across the four HSQ scales, which were mostly low (range from .32–.62, *Mdn* = .40).

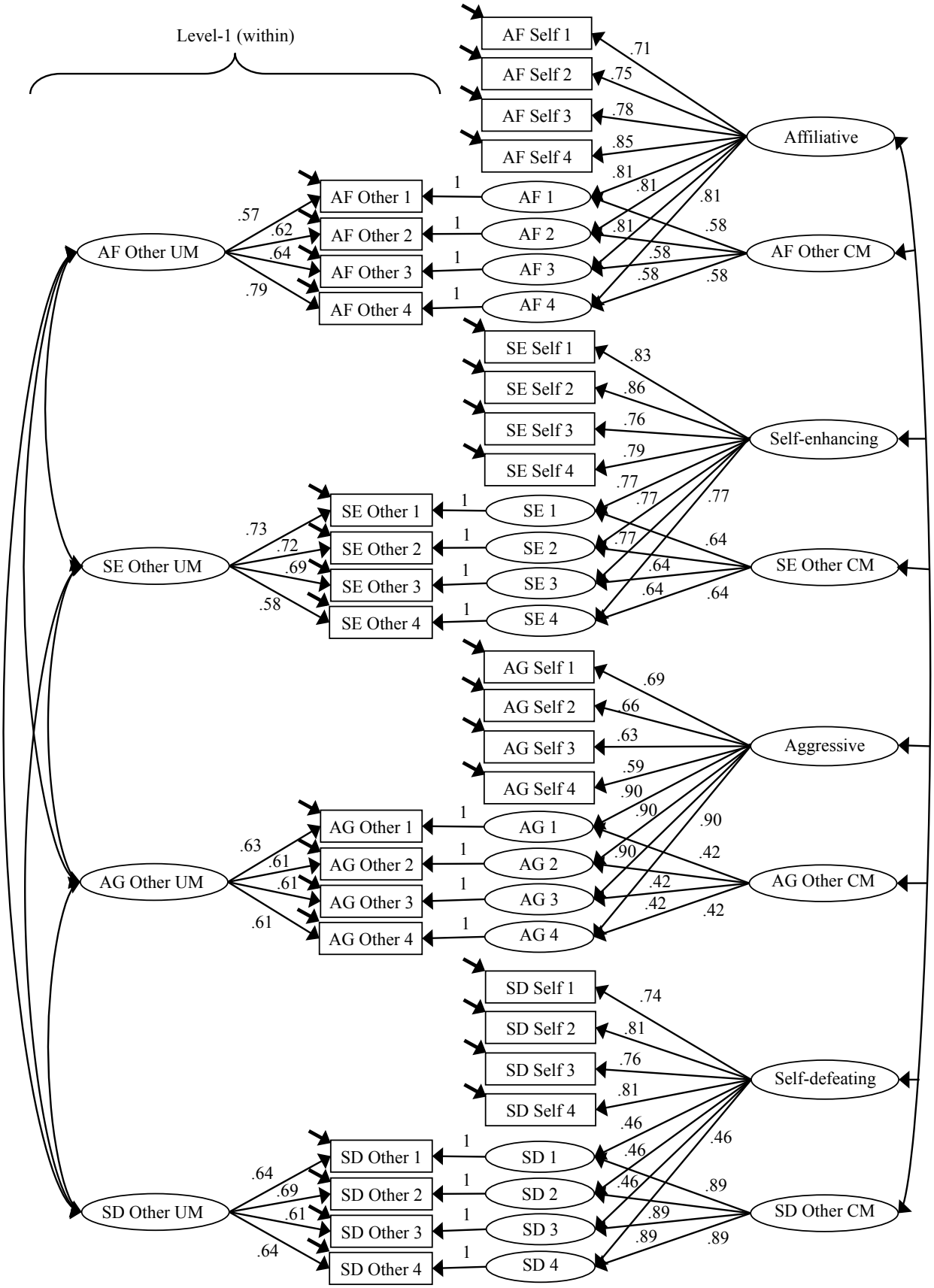

Figure 1. Multilevel multiple-indicator correlated-trait-correlated (method-1) model with standardized loadings (CM = common method, UM = unique method).

Using the formulas by Eid et al. (2008; see Carretero Dios et al., 2011, for another exemplification of these formulas in the context of the ML-CT-C[M-1] model), the observed other-reports can be separated into different coefficients based on the four variance components: Consistency (convergent validities between self-reports and other-reports that consider the average other-report and the individual view of each rater), common method specificity (common view of both raters that is not shared with the target), unique method specificity (individual view of each rater that is not shared with the target and the other rater), and reliability (amount of true score variance in comparison to observed variance). For each HSQ scale, the largest proportion of variance was due to the individual view of each rater (unique method specificity): 49% (affiliative), 57% (self-defeating), 65% (aggressive), and 69% (self-enhancing). The second largest proportion of the variance was explained by the consistencies (convergent validities) for affiliative (33%), self-enhancing (19%), and aggressive (28%). The third largest amount of variance was explained by the common view of the raters (common method specificity): 18% for affiliative, 13% for self-enhancing, and 7% for aggressive. For self-defeating, the variance due to the common method specificity (34%) was larger than the variance due to the consistency (9%). The level-2 reliabilities of the observed indicators of the other-reports ranged from .47–.77 (*Mdn* = .53).

Discriminant validities are indexed by the intercorrelations of the trait factors in the ML-CT-C(M-1) model. (Variances and correlations of the trait and method factors are shown in Table 7 in the Appendix.) The overlap ranged from 0.4% (self-enhancing and aggressive) to 22% (affiliative and self-enhancing), with a median of 6%. The second largest overlap emerged between self-enhancing and self-defeating (11%). The latent correlations among the unique and common method factors additionally showed that method effects did not generalize across each trait.

Overlap with Maladaptive Personality

Correlations and partial correlations (controlling for gender and age) were computed between the self- and other-reported HSQ scales and the self-reported maladaptive personality factors. (Table 8 in the Appendix shows each zero-order and partial correlation.) In line with previous findings of the self-reported HSQ, affiliative (r_{mean} = -.12, $r_{p\ mean}$ = -.14) and self-enhancing (-.17/-15) correlated on average negatively with maladaptive personality, while aggressive (.22/.19) and self-defeating (.19/.19) correlated positively with it. Affiliative (-.07/-.12) and self-enhancing (-.18/-.19) remained adaptive in the other-reports, and aggressive (.10/.07) remained maladaptive. The correlations with self-defeating (-.03/-.02) were close to zero in the other-reports.

Discussion

The present study aimed at testing the self-other agreement of the four HSQ scales, yielding information on the degree of their construct validity. Discriminant validities were high for all HSQ scales, with the largest overlap occurring between affiliative and self-enhancing (22.1% shared true-score variance), which is similar to the previous findings (Martin et al., 2003; Heintz & Ruch, 2015; Zeigler-Hill et al., 2013). The other-reported HSQ scales were mainly determined by unique method specificity (i.e., the individual view of each rater). The second largest source of variance in the other-reported affiliative, self-enhancing, and aggressive scales were the consistencies (i.e., the convergent validities of the self- and other-reports), followed by the common method specificity (i.e., the common view of the raters not shared with the target). Obtaining larger consistencies than common method specificities (ratio 1.5–4:1) supports the convergent validities of these three HSQ scales. For the self-defeating scale, the common method variance exceeded the consistencies (ratio 3.8:1). Similar convergent validities were found when considering the latent correlations at level-2: The amount of variance explained by the self-reports in the average other-reports was large, and larger than the residual for three of the HSQ scales (ratio 1.4–4.6:1), and this effect was reversed for self-

defeating (ratio 3.7:1). Thus, the other-reported HSQ self-defeating scale was more strongly determined by method than by trait variance, failing to support convergent validity for this scale.

Furthermore, the relationships of the self-reported HSQ scales with maladaptive personality were replicated in the other-reports for all scales except for self-defeating (which was found to be neutral in the other-reports). Thus, the relationships to external criteria generalized across other-reports for three of the four scales, further supporting their social reality.

What are possible explanations for the low convergent validity and social reality of the self-defeating scale? As indicated by the large common method specificity found in the ML-CT-C(M-1) model and large observed inter-rater agreements, the two raters agreed on their judgments of the target's self-defeating score more than they agreed with the target's perspective. Stated differently, both raters differed systematically from the target's perspective. This finding reminds of Martin et al.'s (2003) assertion that "Although individuals who are high on this humor dimension may be seen as quite witty or amusing […], there is an element of emotional neediness, avoidance, and low self-esteem underlying their use of humor" (p. 54). According to this view, the raters might have put more emphasis on the observable humor behaviors incorporated in this scale (like letting others laugh at oneself and saying funny things that put oneself down), while they might have missed (or put less emphasis on) the underlying negativity than the targets did (like getting carried away or going overboard when showing these behaviors).

However, given that raters were very familiar with the targets, a more likely explanation might be that the construct validity and social reality of the self-defeating scale is indeed impaired. This implies that the self-defeating scale does not adequately measure the self-defeating humor style and might thus lead to misleading conclusions (see also Ruch & Heintz, 2017). It seems likely that the self-defeating scale is less a measure of humor than of

negative self-evaluation, potentially changing the meaning of the construct and the interpretations of existing findings on the scale. In other words, low self-esteem would not only underlie this humor style (as suggested by Martin et al., 2003), but it might actually be the active ingredient that causes the previously established negative correlations between the self-defeating scale and psychosocial well-being. The other-reports obtained in the present study might have been less influenced by this bias: They showed that the self-defeating scale was not negative in terms of maladaptive personality and also positively related to the self-enhancing scale. The same results were found in studies that focused on the humor entailed in this scale (Ruch & Heintz, 2013, 2017).

Limitations

First, investigating the self-other agreement and construct validity of the HSQ in other languages, cultures, and samples with different demographic backgrounds would be advisable to test to which extent the present findings can be generalized. Second, level-2 reliabilities were lower than .60 for the indicators of the aggressive scale, and level-1 reliabilities were low in general. As the ML-CT-C(M-1) model separates this error variance from the trait and method variance, these lower reliabilities have likely not influenced the present findings substantially. Still, using indicators that consist of more than two items would be desirable in future studies to enhance reliability (as for example shown in Carretero Dios et al., 2011; Koch et al., 2015). Similarly, using at least three instead of two raters per target would be desirable. Third, the present study is purely correlational, and future studies need to go beyond correlations to study the cause and effect in the relationships between the HSQ and psychosocial well-being (e.g., by employing experimental designs or survey testing techniques like cognitive interviews).

Suggestions for Future Research

The present findings highlight the importance of comprehensively investigating the construct validity of humor measures, best during the process of test construction. Testing whether a measure and/or a construct needs further revision at an early stage (i.e., before a measure becomes widely used in research and/or application) would likely avoid interpretation problems. If problems are detected at a later stage (as is the case for the self-defeating scale), either the scale or the construct or both need to be revised, which is a task for future studies.

Employing the ML-CT-C(M-1) model yields information about the extent with which different assessment methods converge with the "golden standard" and which sources of variance contribute to non-convergence. Misalignments can then be resolved by specifically adapting the relevant aspects of the construct and/or measure. Additionally incorporating relevant criteria can show whether the supposed nomological network generalizes across different methods. Overall, the ML-CT-C(M-1) model can be recommended as a useful framework for conducting multitrait-multimethod analyses for psychometric measures if (a) structurally different and interchangeable methods are investigated, (b) a reference method (i.e., a golden standard) can be rationally determined, and (c) some methods are nested within other methods. Recent extensions of the ML-CT-C(M-1) model also allow incorporating structurally different methods at level-1 (Koch, Schultze, Burrus, Roberts, & Eid, 2015) and longitudinal designs (Koch, Schultze, Eid, & Geiser, 2014), making the model suitable for wide range of methods (e.g., different types of raters, observable behaviors such as facial displays of emotion and laughter, or experience sampling data).

Summary and Conclusions

The present study investigated the self-other agreement of the HSQ as an aspect of construct validity, using a large sample of targets and raters (two knowledge informants). Employing the ML-CT-C(M-1) model allowed separating the different variance components entailed in

the multitrait-multimethod data. Convergent validity was supported and correlations with external criteria (maladaptive personality) were replicated for all HSQ scales except for self-defeating, while discriminant validity was sufficient for all HSQ scales. The current findings thus suggest that the self-defeating scale might not measure what is supposed to measure (i.e., a maladaptive humor style), potentially changing the interpretation of past findings and cautioning against a further usage of this scale.

Acknowledgments

I would like to thank Claudia Hürzeler for her valuable help in constructing and pre-testing the other-report form of the HSQ, in setting up the online study, and in recruiting and managing participants. I would also like to thank Willibald Ruch for his feedback on the study design and on earlier versions of this manuscript, Hugo Carretero Dios for his advice on the ML-CT-C(M-1) model, and Nadine Eggimann for her comments on earlier versions of this manuscript.

Appendix: Supplementary Tables

Table 5
Means, Standard Deviations, Inter-Rater-Agreement, and Self-Other Agreement of the Self- and Other-Reports of the Items and Scales of the Humor Styles Questionnaire

	Self-reports		Other-reports		Inter-rater agreement		Self-other agreement	
HSQ	*M*	*SD*	*M*	*SD*	$ICC_{1,2}$	95% *CI*	r_{SO}	95% *CI*
AF	43.42	7.34	42.77	6.55	.63***	[.51, .72]	.59***	[.48, .68]
Item 1	5.63	1.32	5.39	1.21	.53***	[.38, .64]	.47***	[.34, .59]
Item 5	4.56	1.27	4.98	1.14	.42***	[.23, .56]	.31***	[.15, .44]
Item 9	4.75	1.46	4.49	1.15	.24*	[-.01, .42]	.17*	[.02, .30]
Item 13	6.10	1.00	6.04	0.89	.53***	[.39, .65]	.48***	[.34, .59]
Item 17	5.34	1.48	5.15	1.25	.44***	[.27, .58]	.38***	[.26, .50]
Item 21	5.84	1.07	5.38	0.99	.41***	[.22, .55]	.46***	[.34, .56]
Item 25	5.92	1.37	5.78	1.09	.40***	[.21, .54]	.47***	[.35, .59]
Item 29	5.28	1.29	5.57	0.99	.44***	[.26, .57]	.40***	[.27, .52]
SE	35.11	8.79	33.74	6.19	.45***	[.27, .58]	.48***	[.36, .58]
Item 2	4.39	1.51	4.13	1.06	.05	[-.24, .29]	.26***	[.12, .38]
Item 6	4.94	1.46	4.58	1.07	.32**	[.11, .49]	.28***	[.14, .42]
Item 10	3.78	1.53	3.74	1.02	.13	[-.15, .34]	.25***	[.13, .36]
Item 14	4.36	1.54	4.31	1.24	.52***	[.36, .63]	.48***	[.37, .59]
Item 18	3.76	1.62	3.88	1.04	.39***	[.19, .53]	.33***	[.20, .45]
Item 22	4.01	1.61	3.72	1.15	.32**	[.10, .48]	.31***	[.19, .42]
Item 26	4.83	1.30	4.35	1.08	.37***	[.17, .53]	.21**	[.07, .35]
Item 30	5.05	1.39	5.03	0.98	.28**	[.05, .46]	.38***	[.26, .48]
AG	28.86	7.23	27.37	6.75	.44***	[.26, .58]	.52***	[.43, .60]
Item 3	3.35	1.53	3.18	1.32	.40***	[.20, .54]	.32***	[.19, .45]
Item 7	3.46	1.36	3.19	1.25	.32**	[.11, .49]	.32***	[.18, .44]
Item 11	4.32	1.44	4.05	1.24	.16	[-.11, .36]	.23**	[.11, .34]
Item 15	3.12	1.65	3.25	1.25	.13	[-.15, .34]	.31***	[.17, .45]
Item 19	4.50	1.42	3.95	1.29	.29*	[.06, .46]	.28***	[.15, .41]
Item 23	3.60	1.54	3.39	1.28	.37***	[.17, .52]	.47***	[.35, .58]
Item 27	2.85	1.58	2.85	1.26	.38***	[.19, .53]	.38***	[.24, .50]
Item 31	3.67	1.65	3.51	1.24	.29**	[.07, .47]	.37***	[.23, .50]
SD	25.81	8.20	22.66	6.00	.50***	[.34, .62]	.31***	[.18, .44]
Item 4	3.00	1.38	2.61	1.10	.40***	[.21, .55]	.22***	[.08, .35]
Item 8	3.02	1.54	2.41	0.99	.25*	[.01, .43]	.28***	[.14, .43]
Item 12	3.82	1.54	3.08	1.15	.37***	[.17, .52]	.24***	[.10, .37]
Item 16	3.74	1.51	3.29	1.14	.32***	[.10, .48]	.21***	[.06, .36]
Item 20	2.68	1.38	2.20	0.94	.49***	[.33, .62]	.26***	[.11, .41]
Item 24	2.89	1.39	2.56	1.03	.27*	[.04, .45]	.28***	[.13, .43]
Item 28	3.16	1.52	3.13	1.19	.27*	[.03, .45]	.08	[-.07, .24]
Item 32	3.50	1.42	3.37	1.17	.38***	[.18, .53]	.21**	[.07, .34]

Note. $N = 202$. $ICC_{1,2}$ = intra-class correlation (oneway random model), 95% *CI* = 95% confidence interval.
* $p < .05$. ** $p < .01$. *** $p < .001$.

Table 6
Means, Unstandardized Factor Loadings, Residual Variances, and Reliabilities of the Multilevel Multiple-Indicator Correlated Trait-Correlated (Method-1) Model

Indicators	*M*	Trait	CM	UM	Residual variance	Reliability
Affiliative						
Self-report						
AF 1	5.04	1.00			0.68	.50
AF 2	5.78	1.00			0.54	.56
AF 3	5.20	1.00			0.44	.61
AF 4	5.69	1.00			0.27	.72
Other-report						
AF 1	4.82	0.61	1.00	1.00	0.81	.49
AF 2	5.59	0.61	1.00	1.00	0.61	.56
AF 3	5.18	0.61	1.00	1.00	0.56	.59
AF 4	5.80	0.61	1.00	1.00	0.23	.77
Self-enhancing						
Self-report						
SE 1	4.41	1.00			0.50	.69
SE 2	4.36	1.00			0.38	.74
SE 3	4.20	1.00			0.78	.58
SE 4	4.59	1.00			0.64	.62
Other-report						
SE 1	4.45	0.37	1.00	1.00	0.41	.66
SE 2	4.16	0.37	1.00	1.00	0.41	.66
SE 3	3.93	0.37	1.00	1.00	0.91	.47
SE 4	4.33	0.37	1.00	1.00	0.73	.52

(Table 6 is continued)

Table 6 (continued)

Indicators	*M*	Trait	CM	UM	Residual variance	Reliability
Aggressive						
Self-report						
AG 1	3.96	1.00			0.66	.48
AG 2	4.08	1.00			0.78	.44
AG 3	3.29	1.00			0.89	.40
AG 4	3.10	1.00			1.12	.35
Other-report						
AG 1	3.75	0.65	1.00	1.00	0.85	.51
AG 2	3.70	0.65	1.00	1.00	0.95	.48
AG 3	3.22	0.65	1.00	1.00	0.95	.48
AG 4	3.01	0.65	1.00	1.00	0.97	.48
Self-defeating						
Self-report						
SD 1	3.09	1.00			0.73	.55
SD 2	2.79	1.00			0.46	.66
SD 3	3.66	1.00			0.67	.58
SD 4	3.37	1.00			0.47	.66
Other-report						
SD 1	2.76	0.26	1.00	1.00	0.60	.54
SD 2	2.37	0.26	1.00	1.00	0.47	.60
SD 3	3.22	0.26	1.00	1.00	0.69	.50
SD 4	2.94	0.26	1.00	1.00	0.61	.53

Note. AF = affiliative, SE = self-enhancing, AG = aggressive, SD = self-defeating, Trait = trait factor loading, CM = common method factor loading, UM = unique method factor loading. Estimates refer to models in which affiliative was always included. Results only differed slightly in the other models.

Table 7
Variances and Correlations of the Trait and Method Factors in the Multilevel Multiple-Indicator Correlated Trait-Correlated (Method-1) Model

Factor	1	2	3	4	5	6	7	8	9	10	11	12
1. Affiliative	*.70*											
2. Self-enhancing	.47	*1.08*										
3. Aggressive	.22	.06	*.60*									
4. Self-defeating	.17	.33	.27	*.90*								
5. AF Other CM	–	.04	.12	-.17	*.14*							
6. SE Other CM	.18	–	-.19	-.37	-.07	*.10*						
7. AG Other CM	.30	.30	–	-.64	.46	1.28	*.06*					
8. SD Other CM	-.07	.15	-.35	–	.49	.44	.86	*.22*				
9. AF Other UM	–	–	–	–	–	–	–	–	*.39*			
10. SE Other UM	–	–	–	–	–	–	–	–	.54	*.55*		
11. AG Other UM	–	–	–	–	–	–	–	–	.07	.00	*.57*	
12. SD Other UM	–	–	–	–	–	–	–	–	-.03	.11	.30	*.42*

Note. AF = affiliative, SE = self-enhancing, AG = aggressive, SD = self-defeating, CM = common method, UM = unique method. Variances are shown in italics. Estimates refer to models in which affiliative was always included. Results only differed slightly in the other models.

Table 8

Zero-Order Correlations and Partial Correlations (Controlling for Gender and Age) of the Self- and Other-Reported Scales of the Humor Styles Questionnaire (HSQ) with the Five Maladaptive Personality Dimensions

HSQ	NA	Detachment	Antagonism	Disinhibition	Psychoticism
Zero-order					
Self					
AF	-.15*	-.38***	.03	.079	-.16*
SE	-.37***	-.32***	-.07	.03	-.08
AG	-.00	.11	.50***	.31***	.15*
SD	.20**	.15*	.16*	.17*	.26***
Other					
AF	-.09	-.29***	.00	.08	-.03
SE	-.33***	-.35***	-.12	-.03	-.06
AG	-.09	.02	.34***	.24**	-.04
SD	.05	-.05	-.12	-.03	.00
Partial[a]					
Self					
AF	-.20**	-.40***	.02	.07	-.19**
SE	-.35***	-.31***	-.06	.05	-.05
AG	-.01	.07	.47***	.28***	.10
SD	.20**	.15*	.16*	.16*	.26***
Other					
AF	-.15*	-.34***	-.03	.03	-.09
SE	-.32***	-.37***	-.14*	-.03	-.05
AG	-.09	-.02	.30***	.21**	-.08
SD	.12	-.06	-.15*	-.01	.02

Note. $N = 202$. AF = affiliative, SE = self-enhancing, AG = aggressive, SD = self-defeating, NA = negative affectivity.

* $p < .05$. ** $p < .01$. *** $p < .001$. (two-tailed)

[a] Significant gender differences occurred for aggressive (self- and other-report), negative affectivity, and antagonism; significant age differences occurred for aggressive (self-report), affiliative (other-report), negative affectivity, disinhibition, and psychoticism.

PART II: Putting a Spotlight on Daily Humor Behaviors: Dimensionality and Relationships with Personality, Subjective Well-Being, and Humor Styles

Sonja Heintz

This part is an adapted version of the following article:

Heintz, S. (2017). Putting a spotlight on daily humor behaviors: Dimensionality and relationships with personality, subjective well-being, and humor styles. *Personality and Individual Differences, 104*, 407–412. doi:10.1016/j.paid.2016.08.042

Abstract

Although psychological humor research has expanded in the last decades, the humor behaviors that people show in their everyday lives are still poorly understood. To fill this gap, this diary study explores the dimensionality of 45 daily humor behaviors and their relationships with the Big Five personality traits and subjective well-being. Furthermore, the humor behaviors were utilized to investigate the criterion validity of the Humor Styles Questionnaire. A hierarchical factor analysis of the humor behaviors ($N = 123$) revealed seven dimensions: Cheerful, witty, deriding, amused, sarcastic, self-directed, and canned. These humor behavior dimensions correlated with emotional stability, extraversion, lower agreeableness, and culture/openness. Also cheerful, amused, and self-directed humor behaviors correlated positively with subjective well-being, even when personality and the humor styles were controlled for. The criterion correlations of the humor styles to their constituting humor behaviors were medium to large for affiliative and self-enhancing, and small to medium for aggressive and self-defeating. Overall, investigating humor behaviors seems a promising venue for future research and applications of individual differences in humor.

Keywords: humor behaviors; personality; Big Five; subjective well-being; Humor Styles Questionnaire; criterion validity; daily diaries

Introduction

Individual differences in humor can be measured with various approaches, with self-report questionnaires and tests being most prevalent. By contrast, we know little about the actual humor behaviors that people show in their day-to-day lives. This research gap is important to be filled, however, as individual differences in humor need to have everyday behavioral consequences to be relevant for people`s lives (as Furr, 2009, argued for personality psychology in general). For example, if someone endorses an aggressive humor style, one would expect this person to show more corresponding behaviors (like laughing at, making fun of, and teasing others) on a daily basis than a person scoring lower in this humor style. Understanding individual differences in daily humor behaviors and their relationship to personality and well-being is vital for two reasons: First, it allows disentangling the role of humor and additional elements that are often present in trait-based humor questionnaires (e.g., situations, evaluations, attitudes, functions, motives) and thereby allows a more direct test of the role that humor itself plays in personality and SWB. Second, it can help to develop and improve humor-based interventions and trainings by highlighting the humor behaviors that should best be practiced (or maybe decreased) to potentially enhance SWB.

Thus the present paper uses a longitudinal daily diary design (five consecutive days) to explore the dimensionality of 45 different humor behaviors as well as their relationships to the Big Five personality traits and subjective well-being (SWB), which represents the hedonic side of well-being. Additionally, assessing humor behaviors allows testing the criterion validity of humor measures. This is exemplified with the Humor Styles Questionnaire (HSQ; Martin, Puhlik-Doris, Larsen, Gray, & Weir, 2003) by investigating to what extent the HSQ can predict the humor behaviors it entails.

Everyday Humor Behaviors

A frequently employed measure of individual differences in humor, especially everyday functions of humor related to psychosocial well-being, is the HSQ (Martin et al., 2003). It measures four trait-

like humor styles (Martin et al., 2003): Affiliative (enhancing one`s relationships with others while being benign to oneself), self-enhancing (enhancing oneself while being benign to others), aggressive (enhancing oneself while being detrimental to others), and self-defeating (enhancing one`s relationships with others while being detrimental to oneself). The first two are considered to be adaptive to psychosocial well-being, while the two latter ones are considered to be potentially maladaptive.

Notably, two diary studies assessed everyday humor styles with adapted items from the HSQ in two specific contexts, namely work (Guenter, Schreurs, Van Emmerik, Gijsbers, & Van Iterson, 2013) and romantic relationships (Caird & Martin, 2014). However, these only partly measured actual humorous behaviors, as the HSQ items entail further elements that are neither related to humor nor to behavior (see Ruch & Heintz, 2013, for a more detailed discussion). Also the overlap of the HSQ with these diary assessments (i.e., criterion validity) was not tested.

As this is–to our knowledge–the first diary study assessing humor behaviors, it was aimed to explore a broad and varied, albeit not all-encompassing, sample of humor behaviors. The humor behaviors were comprehensively taken from the HSQ to allow testing its criterion validity. Two additional sources were included to add further humor behaviors that were not covered in the HSQ (e.g., relating to irony and satire or reflective and earthy conduct). The first were the ten styles of humorous conduct presented by Craik, Lampert, and Nelson (1996), which are aligned along five bipolar dimensions (socially warm vs. cold, reflective vs. boorish, competent vs. inept, earthy vs. restrained, and benign vs. mean-spirited). Second, Schmidt-Hidding (1963) proposed the eight comic styles of fun, (benevolent) humor, nonsense, wit, irony, satire, sarcasm, and cynicism. It is first of interest how many dimensions underlie the 45 different humor behaviors derived from these three sources using a factor-analytic approach, which can be interpreted as the minimum amount of dimensions needed to describe individual differences in everyday humor behaviors.

Research Question 1 (RQ1): How many dimensions underlie the 45 daily humor behaviors?

Humor, Personality, and Subjective Well-being

Previous research has frequently studied the relationship between individual differences in humor and broad personality traits. The HSQ usually correlated with the Big Five personality traits in a small to large range (for a meta-analysis, see Mendiburo-Seguel, Páez, & Martínez-Sánchez, 2015). Thus some significant correlations might also emerge between personality and the daily humor behavior dimensions, indicating who tends to show which kinds of humor behaviors in their everyday lives.

RQ2: How do the humor behaviors relate to the Big Five personality traits?

Besides personality, research on humor and SWB has been very active throughout the last decades. Consistent relationships were found especially with the HSQ (e.g., Kuiper, 2014; Martin et al., 2003), supporting the notion that some humor styles might be more adaptive (affiliative and self-enhancing) and others more maladaptive (mainly self-defeating) in terms of SWB. Thus the question arises if the daily humor behaviors are also related to SWB and what the direction of these relationships is.

RQ3: How do the humor behaviors relate to subjective well-being?

Given that both personality (for an overview, see Stones, Worobetz, & Brink, 2011) and the HSQ relate to SWB, it is of interest to test if any relationship of the humor behaviors with SWB remains once personality and the HSQ are controlled for. If so, the frequencies with which people exhibit humor behaviors on a day-to-day basis would capture information relevant for SWB that goes beyond the broad personality dimensions and the specific humor styles. This would further highlight their relevance of considering humor behaviors in future research and applications of humor and SWB.

RQ4: Can the humor behaviors explain variance in subjective well-being over and above the Big Five personality traits and the four HSQ scales?

Criterion Validity of the Humor Styles Questionnaire

Despite its popularity, the psychometric properties of the HSQ, specifically its validity, remain understudied. While the criterion validity of the HSQ in terms of psychosocial well-being received support in many studies (e.g., Kuiper, 2014; Martin et al., 2003), its criterion validity in terms of everyday humor behaviors is unknown. As all humor behaviors entailed in the HSQ were entailed in the diary assessment, they lend themselves as external criteria against which the HSQ can be tested; that is, the scores of the HSQ scales should predict the humor behaviors entailed in them. At least medium-sized, but no perfect overlaps are expected to support criterion validity, as the humor behaviors represent one of two central aspects entailed in the HSQ (humor and its specific functions/uses).

RQ5: To what extent can the four HSQ scales predict their everyday humor behaviors?

Materials and Methods

Participants

Overall 123 German-speaking participants filled in at least three of the five daily diaries in time (i.e., on the same evening). The median age of the sample was 24.00 years (M = 27.68, SD = 10.37) ranging from 18 to 68 years (30.9% men, 69.1% women). Participants were primarily Swiss (65.9%) and German (25.2%). Most participants were well educated, with 47.2% being college/university students, 23.6% having passed tertiary education, 24.4% having A-levels, and 4.8% having < 12 years of education.

Before and after the daily diary assessment, participants were invited to fill in several questionnaires (same questionnaires and same procedure for both Assessment 1 and Assessment 2). The sample of Assessment 1 was the same as for the daily diaries, and 107 participants (31.8% men, 68.2% women) completed Assessment 2. Their median age was 24.00 (M = 27.93, SD = 10.95) ranging from 18 to 68 years.

Humor Behavior Assessment

HSQ Humor Behaviors. The daily HSQ humor behaviors were created by rephrasing the 32 HSQ items. A humor behavior was defined as a single concrete behavior that involved humor and/or laughter, and that can be performed by the participants. These criteria resulted in an overlap of several behaviors across different humor styles; for example, the behavior "Said funny things" was described in the HSQ affiliative, aggressive, and self-defeating items. Also some HSQ items entailed more than one behavior, which was then measured with two or more humor behaviors (e.g., the item "I laugh and joke a lot with my friends" was turned into the behaviors "laughed" and "joked around").

However, four of the HSQ items could not be transformed into humor behaviors, as they did not fulfill the criteria: Two self-enhancing items were too abstract and one aggressive and one self-defeating item did not contain active behaviors that participants could perform. Overall, affiliative was represented with seven behaviors, self-enhancing with six, aggressive with nine, and self-defeating with seven (29 overall). As several behaviors were present in two or more HSQ items, 20 *different* humor behaviors were sufficient to cover the HSQ.

Other humor behaviors. Additionally, 14 humor behaviors were extracted, in a similar fashion as was done for the HSQ, from the five bipolar styles of humorous conduct (1–5 humor behaviors each), and 11 humor behaviors were derived from the eight comic styles (1–3 humor behaviors each). The focus was on selecting humor behaviors that were non-redundant and that supplemented the HSQ humor behaviors. All 45 humor behaviors are listed in Table 11 in the supplementary materials.

Participants indicated the frequency with which they showed these humor behaviors on the present day on a five-point scale. The answer options were pre-tested to capture realistic daily frequencies to allow for sufficient variance in the behaviors: 1 = *Not at all (0 times*), 2 = *rarely (1–2 times)*, 3 = *sometimes (3–5 times)*, 4 = *often (6–10 times)*, and 5 = *very often (11+ times)*.

Questionnaires

Humor Styles Questionnaire (HSQ; Martin et al., 2003; German version by Ruch & Heintz, 2016b). The HSQ consists of 32 items measuring four humor styles (affiliative, self-enhancing, aggressive, and self-defeating) with eight items each. Internal consistency (McDonald`s omega) ranged from .71–.89 and test-retest reliability ranged from .74–.88.

Inventory of Minimally Redundant Scales 25 (MRS-25; Schallberger & Venetz, 1999). The 25-item version of the MRS employs bipolar adjectives to assess the Big Five personality traits extraversion, agreeableness, conscientiousness, emotional stability, and culture (with five items each). McDonald`s omega ranged from .74–.90 and test-retest reliability ranged from .90–.94.

Satisfaction With Life Scale (SWLS; Diener, Emmons, Larsen, & Griffin, 1985). The five-item SWLS measures life satisfaction as the cognitive component of SWB (Diener, Suh, Lucas, & Smith, 1999). McDonald`s omega was .91 for Assessment 1 and .89 for Assessment 2, and test-retest reliability was .93.

Positive And Negative Affect Schedule (PANAS; Watson, Clark, & Tellegen, 1988; German version by Krohne, Egloff, Kohlmann, & Tausch, 1996). The PANAS consists of 20 adjectives, with 10 adjectives each for positive and negative affect, which constitute the affective component of SWB (Diener et al., 1999). McDonald`s omega ranged from .85–.90 and test-retest reliability was .79 and .63 for positive and negative affect, respectively.

Procedure

The data were collected in an online survey. The 45 humor behaviors were presented in a randomized order for each participant and for each day. The daily diaries had to be completed on five consecutive evenings. The humor behavior scores were obtained by replacing the missing values (ranging from 8.9–17.9% for each behavior and missing completely at random as indicated by Little`s MCAR test) using the expectation maximization algorithm, and then averaging the daily scores over the five days. The resulting humor behavior scores were internally consistent, as

indicated by McDonald`s omega ranging from .62 to .90 (median = .84; see Table 11 for more details).

Participants were recruited via several means, including mailing lists, social media platforms, and bulletins at several Swiss universities. They were offered personalized feedback and/or course credit in psychology for their participation. The study was conducted in compliance with the local ethical guidelines.

Data Analysis

The 45 humor behavior scores were subjected to a hierarchical factor analysis (HFA; Goldberg, 2006) to derive their dimensionality (RQ1). This analysis not only arrives at a number of dimensions (as exploratory factor analysis does), but it traces their hierarchical emergence from top down. To this end, a series of principal component analyses are conducted, starting with the extraction of the first unrotated principal component and then subsequent varimax-rotated components until one component emerges on which no humor behavior has its highest loading.

The relationship of the humor behaviors with personality and SWB is investigated by correlating the humor behavior dimensions from the HFA with the MRS-25 (RQ2), SWLS, and PANAS (RQ3). The incremental prediction of the humor behaviors (RQ4) in terms of SWB was tested in hierarchical regression analyses, using gender and age as control variables in the first step, the MRS-25 in the second step, and the HSQ in the third step (both measured at Assessment 1). At step 4, each of the humor behavior dimensions were entered in separate regression analyses to avoid any confounding effects between them. Thus seven regressions each were conducted with life satisfaction, positive and negative affect (measured at Assessment 2) as criteria (21 regressions in total). The criterion validity of the HSQ scales (RQ5) was tested by correlating them with their homologous humor behaviors, which were derived by averaging the humor behavior scores of the six to nine humor behaviors belonging to each HSQ scale (McDonald`s omega ranging from .92–.95).

Results

Dimensionality of the 45 Daily Humor Behaviors

The dimensionality of the humor behaviors (RQ1) was tested in the HFA, which suggested the retention of seven components. The process of the HFA is illustrated in Figure 2.

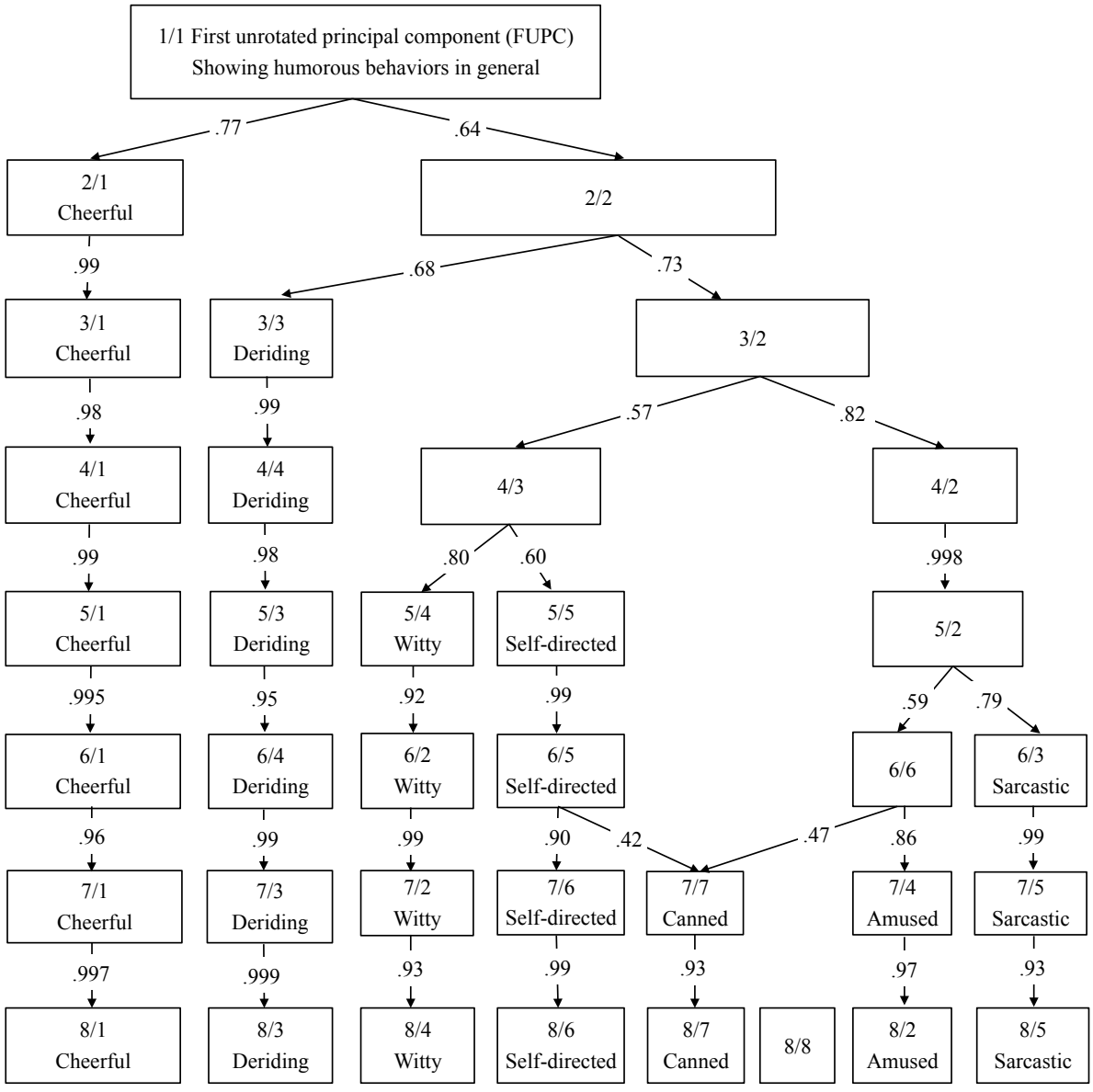

Figure 2. Hierarchical factor analysis of the varimax-rotated components derived from the 45 daily humor behaviors scores.

To analyze the meaning of the humor behavior dimensions, the rotated components matrix of the seven-components solution was investigated (shown in Table 11 in the supplementary

materials). The first component (7/1) was marked by 13 humor behaviors that entailed basic and good-natured humor appreciation (e.g., recognized humorous things) and production (e.g., joked around). Due to its breadth it was labeled *cheerful*, referring to the general propensity to show good-natured humor behaviors. The second component (7/2) was marked by seven humor behaviors that related to wit and sophisticated humor production (e.g., told funny everyday episodes) and was labeled *witty* accordingly.

The third component (7/3) was marked by seven humor behaviors that entailed putting others down and making fun of them (e.g., laughed at someone). The label *deriding* was chosen to indicate a rather blunt and direct mocking of others. The fourth component (7/4) was marked by seven humor behaviors that mostly related to being amused (e.g., amused by absurdities), so this component was labeled *amused*. The fifth component (7/5) was marked by five humor behaviors that involved mockery, cynicism, satire, and irony (e.g., made cynical remarks). The label *sarcastic* was chosen to combine the sophistication of the humor behaviors with the critical component. In comparison to the deriding humor behaviors, the sarcastic ones not only made fun of others, but embedded the criticism in stylistic devices (such as irony) and also aimed at improving the wrongdoings.

The sixth component (7/6) was marked by four humor behaviors that entailed humor directed at oneself (e.g., let someone make fun at my expense), labeled *self-directed*. The seventh component (7/7) was marked by four humor behaviors that mostly related to canned forms of humor (e.g., jokes, nonsense rhymes), and was thus labeled *canned*.

Relationships of the Humor Behaviors with Personality and Subjective Well-being

Having established the dimensionality of the daily humor behaviors, the relationships of these seven dimensions with personality (RQ2) and SWB (RQ3) are of interest, shown in Table 9. (Table 12 in the supplementary materials also shows the correlations of the 45 humor behavior scores with age, sex, personality, SWB, and the HSQ.)

As Table 9 shows, significant relationships occurred between emotional stability and extraversion and cheerful behaviors, between culture and amused behaviors, between (lower) agreeableness and deriding behaviors, and between emotional stability and witty behaviors (small to medium effects). In terms of SWB, the cheerful behaviors positively correlated with higher SWB (medium to large effects). Witty and amused behaviors showed small to medium positive correlations to life satisfaction and positive affect, respectively.

Table 9

Correlations of the Seven Humor Behavior Dimensions with Personality and Subjective Well-being

	Big Five personality traits					Subjective well-being		
Humor behaviors	A	C	ES	E	Culture	LS	PA	NA
7/1 cheerful	.14	.04	.24**	.30***	.09	.31***	.34***	-.35***
7/2 witty	-.04	-.04	.18*	.15	.14	.19*	.17	.01
7/3 deriding	-.19*	-.14	.09	.07	.02	.10	-.08	-.07
7/4 amused	-.02	-.02	.02	-.05	.30***	.10	.22*	.10
7/5 sarcastic	-.11	-.11	.05	-.03	.04	-.11	-.06	.06
7/6 self-directed	.03	.07	.05	.04	-.03	.17	.07	-.01
7/7 canned	-.14	.08	.00	.00	.12	-.06	.07	.03

Note. $N = 123$. A = agreeableness, C = conscientiousness, ES = emotional stability, E = extraversion, LS = life satisfaction, PA = positive affect, NA = negative affect.

* $p < .05$. ** $p < .01$. *** $p < .001$.

Given that some significant correlations emerged, how do the humor behavior dimensions perform in explaining SWB over and above personality and the HSQ (RQ4)? Standard hierarchical regression analyses predicting SWB showed that, consistent with previous studies, personality explained a large amount of variance in the three aspects of SWB ($.41 \leq \Delta R^2 \leq .45$), while the incremental validity of the HSQ was in a medium range ($.05 \leq \Delta R^2 \leq .07$). Importantly, three of the seven humor behavior dimensions explained a small, yet significant amount of variance beyond personality and the HSQ ($.02 \leq \Delta R^2 \leq .04$): The self-directed humor behaviors positively predicted life satisfaction ($\beta = .22$, $p < .01$), and the cheerful ($\beta = .26$, $p < .01$) and amused humor behaviors ($\beta = .17$, $p < .05$) positively predicted positive affect.

Criterion Validity of the HSQ Scales in Predicting Daily Humor Behaviors

The criterion validity of the HSQ scales (RQ5) was tested by correlating them to their homologous humor behavior scores, as presented in Table 10.

Table 10

Correlations of the Humor Styles Questionnaire Humor Behavior Scores with the HSQ Scales

	HSQ humor behaviors scores			
HSQ scales	AF	SE	AG	SD
Affiliative	**.54*** **	.44***	.51***	.48***
Self-enhancing	.29***	**.40*** **	.28**	.26**
Aggressive	.16	.10	**.25** **	.13
Self-defeating	.01	.06	-.05	**.07**

Note. $N = 123$. Criterion validities (diagonal) marked in bold.

** $p < .01$. *** $p < .001$.

Table 10 shows that the criterion validities, that is, the correlations between the corresponding HSQ humor behavior scores and HSQ scales, were medium to large for affiliative and self-enhancing, small to medium for aggressive, and small for self-defeating scale. This indicates that the latter two HSQ scales were less able to predict their corresponding humor behaviors. Interestingly, the HSQ affiliative scale showed medium to large positive correlations to all HSQ humor behavior scores, which were even larger than the criterion validities of the other three HSQ scales. Note that this pattern of correlations was neither due to the overlap of humor behaviors across the HSQ scales nor due to the imperfect mapping of the humor behaviors and the HSQ items. Correlations of the unique humor behaviors of each HSQ style with the HSQ and with a subset of HSQ items that were matched with the unique humor behaviors of each HSQ scale yielded similar results (see Table 13 in the supplementary materials for details).

Discussion

Regarding RQ1, seven higher-order dimensions of humor behaviors emerged in the HFA: Cheerful, witty, deriding, amused, sarcastic, self-directed, and canned. Notably, deriding and sarcastic humor behaviors loaded on separated dimensions, pointing to the importance of distinguishing between blunt and direct versus more sophisticated critical humor. Thus, aggressive or mean-spirited (vs. benign) humor should more appropriately be split into at least two components, and further components might emerge if more humor behaviors from this spectrum (e.g., derived from the comic styles irony, satire, sarcasm, and cynicism) were included. Also some humor behaviors proved to be complex, as they loaded on two dimensions (e.g., “teased someone” had loadings on both cheerful and deriding).

The relationships between the seven humor behavior dimensions and personality (RQ2) supported the notion that the Big Five personality factors (except for conscientiousness) were relevant for four of the seven humor behavior dimensions. Extraversion and emotional stability mainly played a role in humor production (cheerful and witty behaviors), and culture/openness played a role in humor appreciation (amused behaviors). Interestingly, only deriding behaviors

significantly correlated with lower agreeableness while the sarcastic ones were unrelated, which further supports the need to distinguish different aspects of mockery.

In terms of SWB (RQ3), cheerful, witty, and amused humor behaviors positively correlated with one or more aspects of SWB. The cheerful and amused behaviors mostly stemmed from the HSQ affiliative and self-enhancing scales, providing a direct explanation for their relationships with SWB. The witty behaviors, by contrast, were mainly derived from the styles of humorous conduct and the comic styles, and they represent a novel dimension that can be explored in future research on humor and SWB. Importantly, none of the humor behaviors were negatively associated with SWB. This is relevant as the HSQ self-defeating scale was implied to be maladaptive (Martin et al., 2003), which could not be confirmed with the daily humor behaviors, although the self-directed humor behavior component was mainly loaded by humor behaviors from the HSQ self-defeating scale.

RQ4 investigated the incremental power of the humor behaviors over and above personality and the humor styles in predicting SWB. Although this admittedly constituted quite a challenging test, the importance of some of the humor behaviors (cheerful, amused, and self-directed) for predicting life satisfaction and positive affect was supported with small effects. Thus, three of the seven humor behavior dimensions covered unique aspects that are relevant for SWB, which bears a potential for including them in humor trainings and interventions. Curiously, many of the ingredients in McGhee`s (2010) program of training humor habits were reflected in the humor dimensions identified in this study, namely cheerful (general sense of humor, laughter, playfulness), witty (verbal humor), amused (detecting humor in everyday life), and self-directed (laughing at oneself). Also in line with the positive effects elicited by the humor habits program (McGhee, 2010), these four humor behavior dimensions were the ones that were positively related to SWB in the present study.

Finally, the criterion validity of the HSQ scales (RQ5) was supported for affiliative and self-enhancing and to a lesser extent for aggressive and self-defeating. The lower criterion validities of

the HSQ aggressive and self-defeating scales were not caused by general impediments with the prediction of their humor behaviors, as they were all measured reliably and as the HSQ affiliative scale showed medium to large correlations with them. One possible explanation of these findings could be the discrepancy between the humor behaviors and the non-humorous elements (such as evaluations, functions, needs, and motives) entailed in the HSQ scales (see Ruch & Heintz, 2013, for more details). These non-humorous elements (such as going overboard, putting oneself down excessively, or not being able to stop oneself from saying something) might be more relevant to these HSQ scales than the humor entailed in them, thus potentially accounting for the smaller overlap between the scales and their behaviors. These ideas could be tested in future studies.

In sum, the present findings highlight the importance of studying daily humor behaviors, both for theoretical (criterion validity of humor measures) and practical purposes (role of humor behaviors in personality and SWB). The seven dimensions identified in the HFA can be taken as a starting point for future investigations of individual differences in humor behaviors.

Limitations and Suggestions for Future Research

First, the list of humor behaviors was not comprehensive, and future studies should include more humor behaviors to investigate if additional dimensions emerge, also by using additional approaches to generate humor behaviors (like the act frequency approach, as already proposed by Craik & Ware, 2007). Second, the diary methodology employed enabled retrospective daily reports. Employing experience samplings would help to reduce the memory load (and thus potential biases) by prompting participants several times a day to report the humor behaviors they have shown. Additionally, future studies could assess more closely the situations in which the humor behaviors occurred. Third, no causal inferences can be drawn as correlations (albeit in a longitudinal setting) were investigated. Clearly, experimental evidence is necessary to understand the role that humor plays in SWB. Fourth, the sample was not representative of the general population, so replications with more balanced samples and in other cultures and languages are desirable.

Summary and Conclusions

The present paper adds to previous research by suggesting that at least seven dimensions are needed to describe individual differences in daily humor behaviors. Some of these humor behaviors overlapped with the Big Five personality traits and were adaptive in terms of SWB. Importantly, they could uniquely predict SWB once personality and the humor styles were controlled for. The criterion validity of the HSQ was supported for affiliative and self-enhancing, and to a lesser extent for aggressive and self-defeating. Overall, investigating individual differences in everyday humor behaviors can be meaningfully employed to validate humor measures, to better understand their impact on our daily experiences, and to develop and extend humor trainings and interventions to potentially foster people`s life satisfaction and positive experiences.

Acknowledgments

I would like to thank Willibald Ruch for his valuable feedback on the study design and Jenny Hofmann for her helpful comments on earlier versions of the manuscript.

Appendix: Supplementary Tables

Table 11

Description, Source of Origin, Means, Standard Deviations, Reliabilities, and Varimax-Rotated Component Matrix of the Averaged 45 Daily Humor Behaviors (Humor Behavior Scores)

Humor behavior description	Source of origin	M^a	*SD*	ω	7/1	7/2	7/3	7/4	7/5	7/6	7/7	h^2
Found something funny	HSQ AG	3.36	0.92	.87	**.86**	.04	.19	.20	.04	.14	.08	.85
Laughed	HSQ AF, SE, AG	3.66	0.92	.88	**.84**	.14	.11	.09	.07	.08	.03	.75
Thought of something funny	HSQ SE, AG	2.92	0.96	.90	**.81**	.15	.15	.29	.07	.13	.14	.83
Amused someone	HSQ AF	2.82	0.85	.84	**.77**	.28	.25	.16	.24	.20	.11	.87
Made someone laugh	HSQ AF, SD	2.98	0.91	.87	**.77**	.33	.22	.09	.26	.19	.06	.87
Said funny things	HSQ AF, AG, SD	2.75	0.88	.85	**.76**	.35	.23	.19	.28	.16	.07	.90
Recognized humorous things	HBQD RF	2.98	0.90	.87	**.76**	.13	.20	.41	-.01	.08	.13	.82
Joked around	HSQ AF, SD	2.68	0.99	.85	**.74**	.30	.16	.09	.33	.20	.04	.82
Kidded around	HSQ SD, AG	2.32	0.85	.80	**.69**	.28	.20	.12	.35	.18	.17	.79
Cheered myself up with humor	HSQ SE	2.43	0.94	.86	**.58**	.26	.17	.44	-.10	.06	.19	.69
Fooled around innocently	Fun	2.34	0.89	.87	**.55**	.30	.22	.20	.35	.38	-.11	.76
Made witty comments	HBQD EA	2.42	0.85	.85	**.54**	.41	.17	.24	.45	.17	.06	.78
Teased someone	HSQ AG	1.77	0.68	.81	**.48**	.35	**.48**	.14	.27	.31	-.05	.77
Told funny everyday episodes	HBQD RF	1.94	0.67	.79	.42	**.72**	.14	.20	.07	.14	.07	.78
Told stories with punchlines	Wit	1.67	0.69	.84	.39	**.70**	.20	.29	.14	.10	.09	.81
Told funny anecdotes	HBQD CO	1.74	0.68	.80	.41	**.70**	.17	.21	.13	.06	.12	.77
Told funny stories of myself	HSQ AF	1.73	0.68	.84	.42	**.61**	.16	.19	.07	.37	.23	.80
Imitated someone in a funny way	HBQD SW	1.47	0.58	.79	.17	**.52**	.47	.05	.10	.23	.19	.61
Made funny comparisons	HBQD RF	1.91	0.76	.84	.27	**.49**	.16	.41	.43	.22	.18	.76
Caricatured misconduct	Satire	1.60	0.60	.80	.19	**.45**	.28	.34	.35	.43	.17	.76
Laughed at someone	HSQ AG	1.28	0.45	.80	.14	.02	**.87**	.08	.12	.15	.12	.83
Made fun of someone	HSQ AG	1.55	0.60	.81	.24	.13	**.86**	.03	.16	.19	.05	.88
Showed *schadenfreude*	Sarcasm	1.32	0.53	.87	.17	.10	**.75**	.18	.21	-.01	.23	.73

(Table 11 Continues)

Table 11 (continued)

Humor behavior description	Source of origin	*M*[a]	*SD*	ω	7/1	7/2	7/3	7/4	7/5	7/6	7/7	h^2
Bantered someone	HBQD BN (–)	1.91	0.77	.85	.39	.20	**.70**	.11	.20	.20	-.11	.78
Made fun of someone's weaknesses	HBQD BN (–)	1.36	0.41	.67	.23	.29	**.67**	.08	.30	.09	.17	.71
Was amused by the absurdities of life	HSQ SE	2.03	0.87	.88	.35	.21	.08	**.77**	.26	.18	.08	.88
Was amused by the incongruities of life	Humor	1.82	0.80	.88	.24	.18	.09	**.73**	.26	.29	.17	.81
Detected something funny in a situation	HSQ SE	2.46	0.84	.86	.59	.19	.17	**.59**	.15	.15	.05	.81
Thought of an amusing aspect of a situation	HSQ SE	2.40	0.84	.87	.55	.26	.09	**.58**	.19	.18	.14	.80
Amused by grotesque/fantastic stories	Nonsense	1.62	0.58	.69	.28	.32	.34	**.52**	.18	-.08	.13	.62
Made play on words	HBQD BN	1.68	0.72	.85	.19	.47	-.01	**.52**	.25	.00	.21	.64
Joked about problems	HBQD RF	2.09	0.78	.82	.37	.27	.07	**.46**	.44	.20	.11	.67
Made cynical remarks	Cynicism	1.81	0.77	.86	.13	.09	.37	.14	**.79**	.00	.12	.83
Showed scathing mockery	Sarcasm	1.42	0.58	.83	.15	-.09	.33	.32	**.72**	.16	.26	.86
Said something that was ironic	Irony	2.56	0.93	.90	.48	.24	.23	.12	**.58**	.03	.09	.69
Made macabre/grotesque jokes	HBQD EA	1.54	0.58	.78	.22	.35	.27	.42	**.57**	-.04	.14	.76
Ridiculed bad morals	Satire	1.40	0.58	.84	.16	.14	.12	.35	**.53**	.37	.20	.64
Let someone make fun at my expense	HSQ SD	1.48	0.57	.80	.24	.02	.26	.18	.04	**.80**	.15	.81
Let someone laugh at me	HSQ SD	1.82	0.67	.81	.32	.29	.20	.11	.07	**.72**	.23	.81
Said something funny about my own weaknesses, blunders, or faults	HSQ SD	1.83	0.69	.87	.28	**.54**	.04	.25	.18	**.54**	.14	.78
Played the clown	HBQD SW	1.37	0.60	.83	.27	.31	.15	.04	.43	**.47**	.30	.68
Told jokes	HSQ AF, AG	1.34	0.53	.82	.22	.28	.11	.07	.21	.16	**.78**	.82
Exchanged the latest jokes	HBQD BN	1.13	0.30	.74	.06	.50	.23	.24	-.06	.03	**.64**	.78
Used scathing humor against superior people	Satire	1.20	0.31	.62	.02	.00	.15	.16	.41	.25	**.63**	.68
Amused by nonsense rhymes	HBQD BN	1.35	0.59	.84	.12	.03	.12	.41	.21	.29	**.51**	.58

Note. $N = 123$. ω = McDonald's omega, 7/1 = cheerful, 7/2 = witty, 7/3 = deriding, 7/4 = amused, 7/5 = sarcastic, 7/6 = self-directed, 7/7 = canned, h^2 = communalities, HSQ = Humor Styles Questionnaire, AF = affiliative, SE = self-enhancing, AG = aggressive, SD = self-defeating, HBQD = Humorous Behavior Q-sort Deck, SW = socially warm vs. cold, RF = reflective vs. boorish, CO = competent vs. inept, EA = earthy vs. restrained, BN = benign vs. mean-spirited.

Highest loadings of each humor behavior score marked in bold. (–) denotes a reversely keyed humor behavior.

[a] Theoretical minimum = 1 [not at all (0 times)], maximum = 5 [very often (11+ times)].

Table 12

Correlations of the 45 Humor Behavior Scores with Age, Sex, Personality, Subjective Well-being, and Humor Styles

			Big Five personality traits					Subjective well-being			Humor Styles Questionnaire			
Humor behavior descriptions	Age	Sex[a]	A	Co	ES	E	Cu	LS	PA	NA	AF	SE	AG	SD
7/1 Found something funny	-.34*	.00	.08	-.01	.22*	.26*	.13	.31*	.28*	-.27*	.45*	.34*	.16	-.06
7/1 Laughed	-.13	.04	.08	.02	.35*	.38*	.15	.34*	.42*	-.29*	.50*	.33*	.13	-.06
7/1 Thought of something funny	-.30*	-.02	-.02	.02	.17	.19*	.11	.28*	.29*	-.20*	.43*	.33*	.15	.00
7/1 Amused someone	-.24*	-.14	.05	-.02	.34*	.32*	.15	.35*	.38*	-.29*	.54*	.28*	.21*	.05
7/1 Made someone laugh	-.24*	-.08	.07	.04	.38*	.37*	.19*	.33*	.38*	-.28*	.57*	.29*	.15	.02
7/1 Said funny things	-.17	-.09	-.02	-.04	.32*	.26*	.22*	.36*	.35*	-.27*	.51*	.29*	.20*	-.05
7/1 Recognized humorous things	-.23*	-.03	.01	-.03	.21*	.14	.21*	.32*	.33*	-.22*	.38*	.36*	.17	-.07
7/1 Joked around	-.16	.01	.06	-.01	.30*	.30*	.21*	.37*	.35*	-.27*	.47*	.24*	.13	.01
7/1 Kidded around	-.16	-.11	-.02	-.07	.26*	.22*	.18*	.32*	.33*	-.22*	.43*	.19*	.16	-.04
7/1 Cheered myself up	-.10	-.02	.04	.07	.20*	.26*	.28*	.32*	.40*	-.19*	.35*	.41*	.12	.07
7/1 Fooled around innocently	-.02	-.05	.08	-.05	.22*	.21*	.17	.30*	.25*	-.19*	.35*	.16	.07	.11
7/1 Made witty comments	-.15	-.05	.02	-.01	.24*	.20*	.28*	.25*	.30*	-.13	.43*	.23*	.19*	.04
7/1, 7/3 Teased someone	-.18*	-.26*	-.08	-.17	.28*	.22*	.06	.31*	.15	-.24*	.35*	.17	.34*	-.13
7/2 Told funny everyday episodes	-.16	.04	-.08	.07	.19*	.27*	.21*	.26*	.29*	-.06	.39*	.22*	.09	.03
7/2 Told stories with punchlines	-.01	-.12	-.07	-.14	.29*	.19*	.23*	.33*	.27*	-.11	.37*	.23*	.20*	-.04
7/2 Told funny anecdotes	-.16	-.06	-.02	-.13	.25*	.26*	.20*	.30*	.30*	-.10	.35*	.18*	.16	-.02
7/2 Told funny stories of myself	-.14	.08	-.05	.02	.15	.27*	.19*	.31*	.27*	-.05	.35*	.18*	.08	.10
7/2 Imitated someone in a funny way	-.19*	-.11	-.05	-.05	.14	.19*	.22*	.19*	.14	-.01	.26*	.11	.15	-.01
7/2 Made funny comparisons	.01	-.23*	-.10	-.01	.28*	.07	.21*	.22*	.26*	-.07	.24*	.21*	.07	.14
7/2 Caricatured misconduct	-.12	-.16	-.10	-.02	.24*	.14	.20*	.26*	.23*	-.03	.27*	.13	.05	.13
7/3 Laughed at someone	-.18*	-.34*	-.15	-.11	.11	.06	.12	.10	.00	-.06	.20*	.04	.34*	-.04
7/3 Made fun of someone	-.19*	-.31*	-.20*	-.10	.14	.15	.08	.17	.00	-.12	.27*	.04	.38*	-.04

(Table 12 Continues)

Table 12 (continued)

Humor behavior descriptions	Age	Sex[a]	Big Five personality traits					Subjective well-being			Humor Styles Questionnaire			
			A	Co	ES	E	Cu	LS	PA	NA	AF	SE	AG	SD
7/3 Showed *schadenfreude*	-.13	-.35*	-.19*	-.12	.17	.13	.08	.14	.08	-.04	.27*	.09	.41*	-.02
7/3 Bantered someone	-.15	-.21*	.00	-.12	.31*	.26*	.09	.34*	.21*	-.24*	.37*	.20*	.20*	-.04
7/3 Made fun of someone's weaknesses	-.20*	-.27*	-.21*	-.08	.18*	.11	.11	.20*	.10	-.14	.22*	.02	.35*	-.10
7/4 Was amused by the absurdities of life	-.03	-.07	-.03	-.08	.15	.12	.34*	.25*	.31*	-.02	.26*	.28*	.03	.16
7/4 Was amused by the incongruities of life	.03	-.10	-.05	-.04	.22*	.06	.31*	.24*	.34*	-.05	.25*	.28*	.03	.13
7/4 Detected something funny in a situation	-.16	.03	.01	-.02	.09	.17	.28*	.27*	.28*	-.07	.35*	.31*	.04	.06
7/4 Thought of an amusing aspect of a situation	-.09	-.10	-.03	-.02	.16	.17	.31*	.22*	.30*	-.07	.30*	.35*	.01	.08
7/4 Amused by grotesque/fantastic stories	-.07	-.15	.02	-.01	.13	.19*	.33*	.21*	.32*	.01	.30*	.23*	.08	.00
7/4 Made play on words	.20*	-.12	-.04	-.07	.22*	.04	.22*	.13	.23*	-.01	.19*	.16	.02	.10
7/4 Joked about problems	.01	-.14	-.10	-.05	.28*	.08	.11	.21*	.25*	-.16	.26*	.15	.14	.07
7/5 Made cynical remarks	-.22*	-.25*	-.18	-.17	.04	.01	.07	-.04	-.08	.04	.18*	-.03	.30*	.12
7/5 Showed scathing mockery	-.15	-.31*	-.20*	-.11	.04	.01	.15	.01	.03	.08	.14	.00	.25*	.20*
7/5 Said something that was ironic	-.27*	-.13	-.02	-.09	.21*	.17	.09	.12	.12	-.17	.36*	.13	.23*	.03
7/5 Made macabre/grotesque jokes	-.08	-.30*	-.13	-.07	.22*	.18*	.27*	.14	.22*	-.05	.29*	.11	.23*	.06
7/5 Ridiculed bad morals	-.14	-.20*	-.08	-.03	.19*	.16	.22*	.16	.30*	.01	.19*	.08	.03	.21*
7/6 Let someone make fun at my expense	-.11	-.16	-.02	.04	.18*	.12	.09	.24*	.18*	-.13	.20*	.16	.00	.16
7/6 Let someone laugh about me	-.07	-.15	-.05	.00	.25*	.17	.03	.25*	.23*	-.15	.22*	.12	.10	.16
7/6, 7/2 Said something funny about my own weaknesses, blunders, or faults	-.10	.11	-.01	.04	.11	.13	.16	.27*	.21*	.01	.24*	.18	-.05	.27*

(Table 12 Continues)

Table 12 (continued)

Humor behavior descriptions	Age	Sex[a]	Big Five personality traits					Subjective well-being			Humor Styles Questionnaire			
			A	Co	ES	E	Cu	LS	PA	NA	AF	SE	AG	SD
7/6 Played the clown	-.13	-.10	-.08	-.08	.05	.07	.21*	.16	.15	.04	.22*	.08	.07	.18*
7/7 Told jokes	-.11	-.14	-.14	.01	.14	.09	.17	.08	.15	-.06	.19*	-.02	-.02	.07
7/7 Exchanged the latest jokes	.05	-.16	-.13	-.04	.20*	.07	.22*	.16	.22*	-.09	.17	.09	.04	-.01
7/7 Used scathing humor against superior people	-.16	-.15	-.16	.05	.06	.06	.18	.03	.11	.06	.13	-.02	.06	.09
7/7 Amused by nonsense rhymes	-.07	-.04	.02	.01	.08	.11	.17	.12	.13	.02	.11	.16	-.06	.09

Note. $N = 123$. 7/1 = cheerful, 7/2 = witty, 7/3 = deriding, 7/4 = amused, 7/5 = sarcastic, 7/6 = self-directed, 7/7 = canned, A = agreeableness, Co = conscientiousness, ES = emotional stability, E = extraversion, Cu = culture, LS = life satisfaction, PA = positive affect, NA = negative affect, AF = affiliative, SE = self-enhancing, AG = aggressive, SD = self-defeating.

[a] 1 = male, 2 = female.

* $p < .05$.

Table 13

Correlations of the Humor Styles Questionnaire Humor (HSQ) Unique Behavior Scores (based on the Unique Behaviors of each HSQ Scale) with the HSQ Scales and the Sum of Matched Items from each HSQ Scale

	HSQ unique humor behavior scores			
	AF	SE	AG	SD
HSQ scales				
Affiliative	**.50*** **	.36***	.42***	.25**
Self-enhancing	.26**	**.38*** **	.22*	.17
Aggressive	.17	.06	**.35*** **	.02
Self-defeating	.08	.10	-.09	**.23***
Sum of matched HSQ items				
Affiliative	**.46*** **	.23*	.38***	.28***
Self-enhancing	.25***	**.38*** **	.18*	.14
Aggressive	.08	-.01	**.28*** **	-.03
Self-defeating	.08	.14	-.11	**.20***

Note. $N = 123$. Criterion validities (diagonal) marked in bold.

* $p < .05$. ** $p < .01$. *** $p < .001$.

Artikel 3 / PART III

Can Self-Defeating Humor Make You Happy? Cognitive Interviews Reveal the Adaptive Side of the Self-Defeating Humor Style

Abstract:

The present set of studies employs two cognitive interviewing techniques (thinking aloud and online cognitive probing) of the scale assessing the self-defeating humor style, aiming at delineating the role that self-defeating humor plays in self-esteem and emotions. The self-defeating humor style comprises humor to enhance one's relationships with others at the expense of oneself, and has often been related to lower well-being. The analyses are based on 392 item responses of a typical sample (Study 1) and 104 item responses of high scorers on the self-defeating scale (Study 2). Content analyses revealed that higher scores on the self-defeating scale went along with humor (Study 1), with higher state self-esteem, with an improvement of one's interpersonal relationships, and with more facial displays of positive emotions (Study 2). Additionally, the more humor was entailed in the item responses, the higher the state self-esteem and the improvement of relationships was and the more positive emotion words were employed. Thus, the humor entailed in the self-defeating humor style seemed rather beneficial both for oneself and others. These findings call for a reevaluation of past findings with this humor style and provide opportunities for future research and applications of humor interventions to improve well-being.

This part was published as:

Heintz, S., & Ruch, W. (2018). Can self-defeating humor make you happy? Cognitive interviews reveal the adaptive side of the self-defeating humor style. *Humor: International Journal of Humor Research.* Advance online publication. doi:10.1515/HUMOR-2017-0089.

GENERAL DISCUSSION

The present dissertation investigated the construct validity of the HSQ and its role in psychosocial well-being. First, the findings are summarized according to the four aims of the dissertation. Second, implications of the findings for both research and applications are discussed. Third, strengths and weaknesses of the dissertation are highlighted. Finally, general conclusions based on the present findings are provided.

Summary of the Findings in Relation to the Aims of the Dissertation

The present dissertation entailed four aims that address relevant gaps in the existing research concerning the construct validity of the HSQ and the role that humor styles play in psychosocial well-being. The aims were tested across three empirical research papers (Parts I–III), comprising four studies.

Aim 1: Testing the Degree of Self-Other Agreement of the HSQ Scales with a Sound Methodology (Part I)

Part I improved on the methodologies of previous studies of the self-other agreement of the HSQ (Cann et al., 2011; Findlay & Jones, 2005; Martin et al., 2003; Zeigler-Hill et al., 2013) by employing a large sample ($N = 202$), two other-reports (by knowledgeable informants), and the multilevel multiple-indicator correlated trait-correlated (method-1) (ML-CT-C[M-1]) model based on structural question modeling. In this model, convergent validity was assessed as the amount of trait variance (consistency) in comparison to method variance (common method specificity) at level 2 of the multilevel model. Convergent validity was supported for the affiliative, self-enhancing, and aggressive scales. Also discriminant validity was supported for all four scales. Thus, the lower self-other agreement and discriminant validity between affiliative and self-enhancing found in the previous studies might have likely been due to methodical issues, such as short scales to measure the other-reports, including only one knowledgeable informant, small sample sizes, and MTMM approaches that were not based on structural equation modeling.

Still, convergent validity was low for the self-defeating scale both in Part I and in the previous studies, making methodological issues in the studies a less likely explanation for the low self-other agreement of this scale. Rather, this points to problems with either the scale or the construct of the self-defeating humor style, which likely underlie the systematic misperceptions between the person's view of their humor style and how close others perceive it. As this humor style is supposed to serve an interpersonal function (i.e., enhancing one's relationships with others), this lack of social reality bears substantial importance. For example, showing this humor style might not achieve its function of enhancing one's relationships (potentially intended by the target) if others perceive it as hostile or displeasing (see also Kuiper, Kirsh, & Leite, 2010; Martin et al., 2003). Another possibility is that people might not consciously aim at enhancing their relationships by engaging in this humor style, yet the nature of self-directed humor might be beneficial for social relationships. In line with this interpretation, Janes and Olson (2015) noted that "self-deprecating humor tends to be more lubricating than abrasive, because it does not make anyone feel defensive" (pp. 274 f.). Aim 2 delineates in more detail how the self-other discrepancies might have impacted on psychosocial well-being.

Aim 2: Assessing the Relationships of the HSQ Scales and Psychosocial Well-Being Beyond Self-Reports (Parts I and III)

Employing different methods in comparing the HSQ scales with psychosocial well-being enables establishing their overlap free from (or at least with reduced) common method variance (see Podsakoff et al., 2003). Part I found that the other-reported HSQ scales (i.e., how close other perceived the target's humor styles) showed similar relationships with maladaptive personality as the self-reported HSQ did. Specifically, correlations were small to medium and negative for affiliative and self-enhancing, and medium and positive for aggressive and self-defeating in the self-reports. The affiliative and aggressive scales were reduced to small effects in the other-reports of the HSQ, while the negative correlations of the self-enhancing scale remained stable. Only the correlations with the other-reported self-defeating scale dropped to zero; that is, the negativity was

only present in the self-reports of the self-defeating scale, but not in the other-reports. Overall, the criterion validity of the HSQ scales in terms of psychosocial well-being could be replicated for the self-enhancing scale, and with smaller effects for the affiliative and aggressive scales.

Criterion validity was however not supported for the self-defeating scale. Thus, people that were perceived to have higher scores in the self-defeating scale did not report higher scores in maladaptive personality. As the maladaptive personality factors also entail detachment, disinhibition, and antagonism, this finding is also relevant for one's close others. Additionally, this might explain the discrepancy between previous findings that related the self-defeating scale to lower social well-being (e.g., Martin et al., 2003) and the supposed positive interpersonal function of this style. Specifically, the negativity entailed in this humor style might be so pervasive that it not only captures a negative self-evaluation, but a negative outlook on life in general (including one's relationships or attitudes towards others). However, people that were perceived by others to engage in this humor style did not actually have higher scores in maladaptive personality. In other words, although those with higher scores in the self-defeating scale judged themselves to be more detached, disinhibited, and antagonistic, they might not actually engage in such behaviors that harm others or that impair their social relationships.

Part III extended the investigations of the self-defeating scale to a mixed-method approach, which entails quantitative and qualitative methods. Specifically, cognitive interviewing techniques were combined with content ratings, word frequency analyses, and facial displays of emotion. Across two studies, the self-defeating scale was not found to be negative, but rather positive in terms of content-rated self-evaluation (self-esteem), word frequency analyses, and facial displays of emotion. This was found for the "typical" scorers of the self-defeating scale (who on average more disagreed than agreed to the items) as well as for high scorers (who on average agreed to the items). Thus, even though there might be a negativity or low self-esteem underlying the self-defeating humor style (as described by Cann & Colette, 2014; Martin et al., 2003), the interpretation of the items of this humor style was not found to be negative. This also fits to the findings by Cann and

Collette (2014), who did not find significant relationships of the self-defeating scale with daily-measured positive and negative affect. Additionally, the interpretations of the high scores in the self-defeating items (i.e., the reasons why participants chose to agree with the items) entailed a positive self-evaluation and they were accompanied by facial displays of positive rather than negative emotions. An alternative interpretation is thus that the humor entailed in the self-defeating humor style might rather buffer against the underlying negativity, which is discussed in Aims 3 and 4.

Aim 3: Delineating the Humor Entailed in the HSQ Scales (Parts II and III)

Part II extracted the humor entailed in the HSQ scales, which resulted in 20 different humor behaviors. Relating participants' frequencies of showing each of these behaviors across three to five days with their self-reported HSQ scores informs about their criterion validity as well as convergent and discriminant validity. First, the affiliative and self-enhancing scales largely converged with their corresponding aggregated humor behaviors, and the aggressive scale showed medium to large overlaps with its humor behaviors. The overlap of the self-defeating scale with its humor behaviors was close to zero and non-significant. Thus, the criterion and convergent validity in terms of the frequency of showing everyday humor behaviors could be supported for the three of the four HSQ scales. The limited criterion and convergent validity of the self-defeating scale extends the previous findings, and it might explain the lower self-other agreement (and social reality) found in Part I. Specifically, the observable humor behaviors might have been salient for close others and thus influenced their judgment, yet at the same time these humor behaviors did not match with the self-defeating scale. This would foster discrepancies between the judgments of self and others as well as between the scale and the daily humor behaviors. The finding also replicates the previous result that the self-defeating scale was only determined to a limited extent by the humor entailed in it (Ruch & Heintz, 2017). Thus, manipulating the HSQ items to only contain humor and measuring daily humor behaviors yielded a similar pattern of results, indicating that the lack of humor entailed in the self-defeating scale is a robust and generalizable finding as it was corroborated across two

independent methods. At the same time, this convergence supports the usefulness of the two methods (rephrasing items and daily diaries) employed.

Furthermore, the affiliative and self-enhancing scale correlated significantly and positively with the humor behaviors entailed in all HSQ scales. These effects were large for the affiliative scale and larger than the corresponding convergent correlations of the other HSQ scales. Similarly, the effects were medium to large for the self-enhancing scale and larger than the convergent validities of the aggressive and self-defeating scales, indicating a lack of discriminant validity. In other words, the HSQ scales were not able to distinguish among the humor behaviors entailed in them. Instead, all humor behaviors entailed in the HSQ were compatible with the affiliative and self-enhancing scales. Additionally, no significant relationships were found between the humor behavior scores and the self-defeating scale, indicating that this scale was incompatible with the humor behaviors in general. This again fits to the lack of humor entailed in the self-defeating scale (Ruch & Heintz, 2017).

Part III found that some of the interpretations of the HSQ self-defeating items entailed self-directed humor (i.e., humor, jokes, laughter, and amusement about oneself). In other words, participants differed in the extent to which they reported humor when responding to the self-defeating scale. However, the means of the self-directed humor reported were below the scale midpoint across both studies, which is in line with the results of Part II and Ruch and Heintz (2017). Importantly, this self-directed humor might not match with the self-defeating scale itself, and it might instead represent a more general type of humor that is compatible with the affiliative and self-enhancing scales (as was found in Part II). This would also fit well with McGhee's (1999, 2010) notion that laughing at yourself (akin to self-directed humor) is a central and positive aspect of one's sense of humor. In other words, the difference between the self-defeating scale and other humor scales seems to be the context that is added to the humor behavior (like going overboard), while the humor behavior itself is similar to that of other scales (such as the affiliative and self-enhancing scales and the laughing at yourself scale from the Sense of Humor Scale; McGhee, 1999,

2010). As this added context is not humorous, primarily negative, and tends to override the variance generated by humor (see Ruch & Heintz, 2017), the resulting scores of the self-defeating scale are biased. Unless specific methods are used to disentangle humor from the other components entailed in the items (by rephrasing items, capturing the humor behaviors, or cognitive interviews), the self-defeating scale cannot be interpreted in terms of humor.

Aim 4: Delineating the Relationship of the Humor Entailed in the HSQ Scales and Psychosocial Well-Being (Parts II and III)

Part II presented the relationships of the seven dimensions based on 45 different daily humor behaviors (of which 20 were taken from the HSQ) with the three components of subjective well-being (life satisfaction, positive and negative affect) and their incremental validity beyond the Big Five personality traits and the HSQ. The cheerful humor behavior factor (which entailed affiliative humor behaviors) showed large positive correlations with subjective well-being, and the amused behavior factor (which entailed self-enhancing humor behaviors) showed medium positive correlations with positive affect. These two humor behavior factors also incrementally explained small to medium (amused) and medium to large (cheerful) variance in positive affect beyond personality and the HSQ. Additionally, the humor behaviors of the self-directed factor (which entailed self-defeating humor behaviors) incrementally and positively explained life satisfaction (medium effect). Although not presented in Part II, additional analyses conducted for the humor behaviors entailed in the HSQ scales replicated these effects (see Table 18).

As shown in Table 18, the daily-measured humor behaviors entailed in the HSQ all correlated positively and strongly with life satisfaction and positive affect. Also medium to large effects were found for the unique humor behaviors entailed in each HSQ scale. Importantly, all HSQ humor behaviors showed medium to large incremental validities in terms of life satisfaction and positive affect. The same was the case for the unique humor behaviors of the HSQ scale, except for the aggressive scale. Thus, the humor behaviors entailed in all HSQ scales could be considered positive in terms of life satisfaction and positive affect, and neutral in terms of negative affect.

Table 18

Correlations and Incremental Validities of the Humor Styles Questionnaire (HSQ) Humor Behavior Scores and Unique Humor Behavior Scores with Subjective Well-Being (Additional Analyses of Data from Part II)

	Zero-order correlations				Betas (incremental validity)			
Subjective well-being	AF	SE	AG	SD	AF	SE	AG	SD
Humor behavior scores								
Life satisfaction	.39***	.35***	.36***	.37***	.25**	.23*	.20*	.27**
Positive affect	.43***	.46***	.39***	.39***	.25**	.31***	.22*	.22*
Negative affect	.00	.01	.00	.00	.15	.08	.14	.11
Unique humor behavior scores								
Life satisfaction	.36***	.29**	.30**	.26**	.27**	.21*	.14	.25**
Positive affect	.39***	.41***	.24*	.27**	.25**	.27**	.10	.18*
Negative affect	.05	.02	.03	.02	.12	.05	.14	.02

Note. $N = 107$. AF = affiliative, SE = self-enhancing, AG = aggressive, SD = self-defeating, Betas = standardized regression weights of a hierarchical regression (entering the HSQ and personality in Step 1, and each of the HSQ humor behavior scores in Step 2). HSQ and personality measured at Assessment 1, subjective well-being measured at Assessment 2, humor behaviors aggregated across 3–5 days.

* $p < .05$. ** $p < .01$. *** $p < .001$.

Furthermore, the incremental validity of the humor behaviors calls into question the comprehensiveness of the HSQ. Specifically, Martin et al. (2003) proposed that the four humor styles should represent most of the elements of the literature on humor and well-being. Establishing incremental validity of humor behaviors beyond the HSQ shows that there is variance in subjective well-being that is not captured by the HSQ scales. Thus, studying humor behaviors offers an added value in the study of humor and subjective well-being, and likely also social well-being. Employing a diverse set of 45 humor behaviors or the 20 behaviors entailed in the HSQ already yielded incremental effects in explaining life satisfaction and positive affect beyond personality and the HSQ. Employing more comprehensive sets of humor behaviors would likely yield even larger incremental effects. Similarly, the HSQ would need to be tested against other humor measures (especially styles of humorous conduct and the comic styles) to show that it is indeed sufficient to explain the variance in psychosocial well-being. Importantly, these relationships would need to control for the context in the HSQ items or for personality, as else the non-humorous elements entailed in the HSQ scales (especially the self-defeating scale) would artificially inflate the explained variance (see Ruch & Heintz, 2013, 2017). Given the similarities of the humor behaviors entailed in the four HSQ scales in terms of subjective well-being, it might however be expected that the HSQ is not sufficient to comprehensively explain the variance in well-being beyond other humor scales.

Part III investigated the relevance of the humor entailed in the self-defeating scale for the affective component of subjective well-being (i.e., positive vs. negative affect) and self-esteem. The amount of self-directed humor entailed in the explanations of the self-defeating items was related to a more positive self-evaluation and to using more positive emotion words. This effect was found with two cognitive interviewing techniques (online cognitive probing and thinking aloud) and across the whole range of the HSQ scale (i.e., low to medium and high scorers) Thus, self-directed humor, elicited by answering the self-defeating scale, was always associated with a more positive self-evaluation (i.e., higher self-esteem) as judged by independent raters and more positive

emotions (as expressed in the words that participants used). Furthermore, the humor entailed in the self-defeating scale also related to improving one's relationships (Study 2).

Implications of the Present Dissertation

The present dissertation has several important implications, both for existing research and applications and for future endeavors in the area. Implications are discussed for research and applications of the HSQ, for the assessment of humor styles and individual differences in humor, and for the role that humor plays in psychosocial well-being. Thus, the present findings can be deemed relevant for many areas in psychology, like personality psychology, psychological assessment, positive psychology, social psychology, emotion psychology, and clinical psychology.

Implications and Directions for Future Research

Affiliative and self-enhancing humor styles and their measurement. The present dissertation supported and at the same time advised caution for research that employed the HSQ. Parts I and II mostly supported the construct validity and relevance for psychosocial well-being for both the affiliative and self-enhancing scales. Although no causal conclusions can be drawn based on the findings, they support the notion that these scales possess a social reality. That is, they were also recognized by close others and they could be distinguished from one another, while the humor behaviors were rather similar. Additionally, the two scales were able to predict the frequency with which participants showed the humor behaviors entailed in them in their day-to-day life. Their humor behaviors also showed unique and positive relationships with life satisfaction and positive affect. Thus, their proposed positive and beneficial nature received further support.

The similarity in their humor behaviors raises the question on which humor aspects actually differentiate between the affiliative and the self-enhancing humor styles. Conceptually, affiliative is more rooted in social interaction and humor production, while the self-enhancing styles is more strongly rooted in a cheerful mood and coping with stress. This is also reflected in the HSQ items, as the affiliative scale includes items set in a social context (i.e., being with friends, family), while

the self-defeating scale emphasizes being alone, facing problems, or being in a bad mood. While this difference in the contexts is apparent, the difference in the humor entailed in the scales is less clear (as indicated by the humor behaviors and by the humor-only version of these HSQ scales; see Ruch & Heintz, 2017). Interestingly, a previous dissertation (Reff, 2006) revised the HSQ by replacing the items by everyday life situations that participants needed to respond to in a humorous way. This revised version found three scales to be superior (positive, aggressive, and self-defeating humor), aggregating the affiliative and self-enhancing scales. Specifically, he noted on the self-enhancing style that "It is difficult to forge situation-response style items based on this humor style that do not elicit the potential positive impact of those around you" (p. 46). Thus, future research is needed to show how the affiliative and self-enhancing humor styles differ from one another, especially in terms of their humor content, and whether they can be separated from existing humor constructs (such as cheerfulness).

Future directions with these two humor styles would include establishing causal effects. As a first step in this direction, Ford et al.'s study (2017) found that engaging in the self-enhancing humor style lead to a lower state anxiety in comparison to engaging in the self-defeating humor style and a non-humor control condition. Future studies should employ experimental manipulations with pre-post designs, in which they investigate different humor-related styles against control conditions and also against other humor constructs (such as humor behaviors). This would help to show (a) which aspects of humor can enhance psychosocial well-being, (b) which kinds of humor are more effective than others, and (c) which variables mediate and moderate the effectiveness of humor interventions (e.g., situations, mood, personality traits, or the sense of humor). Past research already found positive effects on psychological well-being by the seven humor habits program (McGhee, 2010; for an overview, see Ruch & McGhee, 2014) and for online humor-based positive psychology interventions (Wellenzohn et al., 2016a, 2016b; for an overview, see Ruch & Hofmann, 2017). As already established humor interventions exist, the added value of the affiliative and self-enhancing humor styles would need to be tested. In general, delineating the working mechanisms

and the specific humor aspects that make humor interventions more or less effective are fruitful next steps in the research on humor and well-being.

The aggressive humor style and its measurement. The aggressive scale received support for its construct validity in terms of self-other agreement, while its construct validity in terms of everyday humor behaviors was less supported. Also, the role that the aggressive scale plays in psychosocial well-being (as criterion validity) was inconsistent. More specifically, the negativity in terms of maladaptive personality was only found in self-reports and to a lesser extent in other-reports. By contrast, the daily humor behaviors of the aggressive humor style were positive in terms of subjective well-being (especially life satisfaction and positive affect). These findings more closely reflect the concept of the aggressive humor style, which was defined to aim at enhancing oneself, yet neither the predictions put forward by Martin et al. (2003) nor the items of the HSQ aggressive scale entail this positive aim. Overall, this scale should be studied further to determine which effect it actually has on psychosocial well-being. Besides psychological well-being, investigations of the everyday humor behaviors of the aggressive humor style should be extended to social well-being. For example, including reports of people the person interacted with could show if the aggressive humor style is always perceived in a similar fashion (i.e., as criticizing and aggressive), or whether people might also perceive playful or teasing aspects. Also situational factors (like mood, the relationship between the target and the joker, and the type of situation) should be taken into account to discover systematic effects to elucidate how this humor style impacts one's social well-being.

Importantly, Part II showed that the aggressive humor style did not represent all kinds of "aggressive" humor, or, in other words, that the latter is not a unidimensional construct. The 45 daily humor behaviors yielded two dimensions, namely deriding (which entailed aggressive humor behaviors) and sarcastic , which mainly consisted of the comic styles irony, satire, and sarcasm. This also fits to the less than perfect overlap of the aggressive scale with these comic styles (as presented in the section "Overlaps Among the Different Humor Styles"). Thus, future research

should distinguish among different components of aggressive humor, as they likely have quite different implications for psychosocial well-being. For example, Ruch and Heintz (2016a) found that corrective humor, akin to satire, was uniquely related to interpersonal character strengths (i.e., fairness and bravery). This lack of differentiation of the aggressive humor style might also have contributed to the inconsistencies in both the concept and previous research on this style. Separating the different components likely helps to yield a more coherent picture of the different facets of aggressive forms of humor and their relevance to psychosocial well-being. However, the term 'aggressive' might be less suitable as an overall term, as it already includes the emotional tone or attitude with which this humor is delivered. A more appropriate term might instead be 'mockery' (see also Ruch & Heintz, 2016b; Ruch et al., 2018).

The self-defeating humor style and its measurement. The most detailed analyses were conducted with the self-defeating humor style and its scale. Parts I to III converged in finding little support for the construct validity and social reality of the self-defeating scale. Specifically, the self-other agreement was low and the self-defeating scale did not converge with the daily humor behaviors entailed in the scale. This suggests that the self-defeating scale had a limited ability to predict how close others' perceive a person's humor style and how people act in their everyday lives. Also the self-directed humor behaviors and the humor behaviors entailed in the self-defeating scale were not negative in terms of subjective well-being, in contrast to the self-defeating scale, but they were instead positively related to life satisfaction and positive affect. This suggests that the humor entailed in this scale was not negative, but rather positive, and also explained additional variance beyond personality and the HSQ. Similarly, the negativity of the self-defeating scale was not evident in the other-reports in Part I.

Part III was designed to follow up on the findings from Parts I and II and to examine the process that might be underlying these empirical discrepancies. A possible explanation why the self-defeating scale was found to be negative in most previous studies was offered by the cognitive interviewing techniques employed: Participants differed in the degree to which the interpreted self-

directed humor in the self-defeating items. If participants described a lot of self-directed humor in their explanation of the item responses, they were rated to have a positive self-evaluation and used positive emotion words. This conflicts with the concept of the self-defeating humor style, which posits that engaging too much in self-deprecating humor would be harmful. The key difference is that the 'too much' was assessed in Parts II and III as a quantity (i.e., how much self-directed humor the person reported in response to the self-defeating items, or the frequency with which the behaviors entailed in the scale are shown), showing that simply engaging more in self-directed humor is not inherently harmful. Instead, the negative connotation of the self-defeating humor style and scale seems be derived primarily from the subjective judgment that one shows this style 'too much' by going overboard or by doing it more than one should. As the latter taps into neuroticism or negativity more generally (for a demonstration with personality items, see Haigler & Widiger, 2001), this wording is likely the source of the negative bias underlying the self-defeating scale.

In other words, these findings lend support for the notion that there might be an underlying negativity entailed in the self-defeating humor scale (as described by Martin et al., 2003), and this might also drive the negativity found in terms of psychosocial well-being. When humor is added to the equation, however, the connotation shifts to the positive side. This implies that the humor in the self-defeating scale might have the opposite role than was assumed in research on the self-defeating humor style: Engaging in self-directed humor (or 'excessively self-disparaging humor') might not lower one's well-being, but it might be employed by people with a lower self-esteem to buffer against the negativity they experience. That is, the humor entailed in the self-defeating scale might actually serve as a coping mechanism, as was originally proposed for the self-enhancing scale only (Martin et al., 2003).

Several important implications are associated with these findings and interpretations. First, existing research that employed the self-defeating scale needs to be reevaluated to take into account that not the humor, but rather the negativity entailed in the scale drove the negative associations to psychosocial well-being. Thus, 'self-defeating humor style' seems to be a misnomer, as the scale

entails a negative self-defeating aspect in addition to a positive self-directed humor, yet, the negative self-defeating aspect seems to be more pervasive (see also Ruch & Heintz, 2013, 2017). Second, this also offers a straightforward interpretation of Martin et al.'s (2003) statement that "This is also a humor dimension that appears to be largely untapped by previous humor scales" (p. 71). The lower overlap obtained between the self-defeating scale and existing humor measures (see also Cann, Zapata, & Davis, 2009; Ruch & Heintz, 2016b) might have been found because the scale does not primarily measure humor. In other words, these findings might rather reflect a lack of construct validity of the self-defeating scale instead of highlighting the novelty of this humor style.

Overall, the scale in its current version does not adequately, but rather misleadingly, represent self-directed humor, and it does not seem suitable to assess humor or a humor style. This highlights the need for revisions of the HSQ self-defeating scale due to its low construct validity. Presently, the self-defeating scale cannot be recommended for usage, and past findings with the scale need to be treated with caution (in line with the recommendations by Cronbach & Meehl, 1955). Furthermore, the construct of the self-defeating humor style would benefit from revisions as well. For example, it is the only humor style that is not tied to humor theories or models, but is instead loosely based on two observational studies in a clinical and a school setting (Martin et al., 2003). Delineating the theoretical background of this humor style as well as what is meant by 'too much' conceptually would help to reconcile the conflicting empirical findings on this humor style.

General implications. Going beyond the HSQ, the present findings point to the yet unused potentials entailed in self-directed humor and laughing at oneself specifically. They represent positive kinds of humor that should be included in future studies on humor interventions to delineate their causal effects on psychosocial well-being. The present findings suggest that they might have incremental positive effects in terms of psychological well-being, and potentially also in terms of social well-being by enhancing one's relationships with others, loosening up tense moods, and cheering others up. Additionally, the seven humor behaviors dimensions that could be distinguished in Part II point to the need of developing humor models and instruments with more

than four dimensions (as is already the case for the comic styles and the styles of everyday humorous conduct). Future research should aim at more comprehensively delineating and assessing individual differences in humor to arrive at instruments that actually represent all relevant and distinguishable dimensions. The comprehensive definition of humor styles presented in the General Introduction (section "Humor Styles") could be a useful start of this endeavor. Finally, the dissertation highlights the importance of sound scale construction and especially the importance of item wording, as including terms such as 'too much' likely leads to a potentially unwanted negativity in the items (see Haigler & Widiger, 2001). Following guidelines that ensure construct validation in all stages of test construction (see Simms, 2008) should help to avoid developing and using scales with limited construct validity, which likely hinder rather than foster scientific progress in an area.

Implications and Directions for Future Applications

The HSQ was has already been applied in areas such as work (e.g., Guenter, Schreurs, Van Emmerik, Gijsbers, & Van Iterson, 2013; Scheel et al., 2016), relationships (e.g., Caird, & Martin, 2014; Cann et al., 2009), and psychopathology (e.g., Frewen, Brinker, Martin, & Dozois, 2008; Meyer et al., 2017; Tucker et al., 2013; Zeigler-Hill et al., 2016). The findings in these areas were usually supportive of the positive effects of the affiliative and self-enhancing scales and the negative effects of the self-defeating scale. While the former could be corroborated by the present findings, the correlations of the self-defeating scale with negative outcomes might rather be attributed to a general negativity (like self-esteem), and not to the self-directed humor itself. Applied studies would thus be needed that investigate the humor styles at the level of the humor entailed in them. At the very least, the self-directed humor needs to be separated from its underlying negativity to be able to separate cause and outcome when employing the self-defeating scale.

Based on the present findings, applications that foster different kinds of humor would seem helpful, going beyond the broad 'adaptive/maladaptive' distinction introduced by the humor styles. First, self-directed humor would be good candidate for future interventions in many areas, as it

might be beneficial for several aspects of psychosocial well-being. Specifically, positive psychology interventions, strategies of coping with stress, and social competence trainings could be developed that practice self-directed humor. For example, McGhee (1999, 2010) proposed to practice laughing at oneself by making fun of topics that one feels comfortable with, and then step-by-step integrate more sensitive topics. One could first intentionally plan jokes and funny comments about oneself, and then slowly start to with making spontaneous funny remarks about oneself or the current situation. One could also adapt the daily "three funny things" intervention (Wellenzohn et al., 2016a, 2016b) to entail three funny things about oneself or to reframe three serious things in a humorous manner. Second, the different components of mockery could be included in interventions and trainings to educate and foster its virtuous facets (i.e., mocking moral transgressions to improve them; see Ruch & Heintz, 2016a) and to separate them from potential negative facets (like insulting, criticizing, and putting others down). This is in line with the positive psychological perspective that includes humor as one of 24 character strengths (Peterson & Seligman, 2004), highlighting the potentials of humor-based interventions that might not only increase well-being, but also virtues and morality.

In general, applications need to be based on valid research findings, and more research in the area is certainly needed to discern which types of humor can be effectively employed when and how and with whom. It seems clear that humor styles, the sense of humor, and humor in general need to be regarded as multidimensional constructs. However, which aspects of humor are positive and which of them are negative is likely more complex than can be captured by a simple 2 × 2 conceptualization that underlies the humor styles by Martin et al. (2003). Despite the obvious simplicity and potential appeal of this conceptualization, it does not seem suitable for a comprehensive approach to humor styles, whether restricted to the area of humor and well-being or not. Once the humor facets relevant for psychosocial well-being and strategies to enhance them are delineated in more detail, they could be integrated in existing trainings, therapies, and interventions to enhance psychosocial well-being (e.g., in clinical or work settings).

Strengths and Limitations of the Present Dissertation

The strengths and limitations of the present dissertation are discussed separately for Parts I–III, as they are mostly based on the methods employed. Several arguments generalize across all Parts and are discussed first.

Strengths

General strengths of the present dissertation include the multi-method approach employed in all Parts: Self- and other-reports in Part I, self-report questionnaires and daily diaries in Part II, and self-reports, content ratings, word frequency analyses, and facial displays of emotion as well as a mixed-method approach (using cognitive interviewing techniques) in Part III. Although going beyond self-report questionnaires is effortful, the present dissertation highlights the necessity and benefits of using multiple methods to be able to generalize findings and to detect flaws in one method. A second strength is that converging findings could be established across different outcome measures of psychosocial well-being (maladaptive personality, life satisfaction, positive and negative affect, and self-esteem), which increases the generalizability and robustness of the findings.

Strengths of Part I include the usage of a modern MTMM model, the relatively large sample, and the recruitment of two close others. This enables sound conclusions to be drawn from this study in comparison to previous studies, which might have been partially flawed and might thus have underestimated the degree of self-other agreement of three of the four HSQ scales (all except the self-defeating scale). Part I also highlights the advantages and flexibility of the CT-C(M-1) model in analyzing a wide range of MTMM data. Strengths of Part II include a pre-post design and the inclusion of humor behaviors from several measures to arrive at a broader list. This enabled comparing the HSQ scales to their corresponding humor behaviors and also showed that at least seven dimensions of everyday humor behaviors need to be distinguished empirically. Thus, this study is relevant for humor research in general and the role that humor plays in our everyday lives. Strengths of Part III include the novel combination of methods (i.e., content ratings, word frequency

analyses, facial action coding, and cognitive interviews). This shows the potential for future studies to more comprehensively assess the cognitive processes underlying questionnaire responses as well as their individual differences and affective connotations. Additionally, covering the full spectrum of a construct under question (the self-defeating humor style in Part III) is crucial to draw appropriate inferences for those with low, medium, and high scores (especially if any of the groups is under- or overrepresented in 'standard' samples).

Limitations

The present dissertation has several limitations. First, the studies were conducted in one culture and one language, and replications are needed to more confidently generalize the present findings. Regarding Part I, it would have been advantageous to employ more than two raters and to have more reliable parcels to increase the reliability and representativeness of the other-reports. Additionally, maladaptive personality was not measured in other-reports, and thus it is yet unknown how close others perceive the interplay between the target's humor styles and their maladaptive personality. However, it would need to be considered whether the time and effort would be well spent to further investigate the self-other agreement of the HSQ, or whether rather new scales or a revised version of the HSQ should be developed that already establish construct validity during the construction process (also in terms of content validity). Regarding Part II, the humor behaviors employed were comprehensive for the HSQ, but not for the humor behaviors in general. Thus, employing approaches that cover humor comprehensively (similar to the Craik et al.'s [1996] approach) would be necessary to determine how many and which dimensions of humor behaviors can be distinguished from one another. For example, this could be achieved by assembling all humor behaviors discussed in the literature, entailed in existing scales, derived from act-frequency approaches, and from a psycho-lexical approach. Also the sole employment of self-reports potentially biased findings due to common method variance in the predictor, outcome, and daily-diary measures. Lastly, the humor behavior list was newly generated for the study, and it would require a full test construction process to be recommended for usage in future research and

applications. Regarding Part III, both studies employed rather small samples, limiting the generalizability of the obtained effects. Future studies should employ larger and more representative samples to increase power and generalizability of the findings. As the combination of the methods was novel, Part III was mostly exploratory, again emphasizing the need for replication studies. Despite the promising findings, the possibilities and limitations of the different methods should be delineated in future studies. In general, the present dissertation focused on a small subset of well-being constructs, namely self-esteem, subjective well-being, and maladaptive personality. Future studies should incorporate more comprehensive measures and base their hypotheses on more comprehensive well-being theories to refine the relationships between different aspects of the sense of humor and physical, mental, and social aspects of well-being.

Conclusions

Overall, the present dissertation showed that the affiliative and self-enhancing HSQ scales could be corroborated in terms of construct validity and in their relevance for psychosocial well-being. Still, the added value of these two humor styles beyond existing humor concepts still remains to be shown. The aggressive scale was positively related to antagonism and disinhibition in both self- and other-reports, while the humor entailed in the scale was positively related to life satisfaction and positive affect, yielding overall mixed findings in terms of psychosocial well-being. It thus seems fruitful to separate the aggressive humor style into different dimensions, such as deriding and sarcastic elements, which likely differ in their impact on psychosocial well-being. The self-defeating scale seems to have been a ‘shotgun wedding’ between a general negativity and positive self-directed humor, which might have lead to misleading conclusions about this humor style in past research. The present dissertation separated these elements indirectly and directly and could thus support the positivity of the self-directed humor that is entailed in this scale. Thus, this HSQ scale is in need of revisions, and also the humor style underlying the scale would require further elaborations to better understand its theoretical relevance and derivation.

In conclusion, the usefulness of the HSQ is questionable, as the affiliative and self-enhancing scales strongly overlap with other humor-related styles (e.g., the social warm and competent styles of humorous conduct, and the comic styles fun and benevolent humor), the aggressive scale should rather be subdivided into different components (such as irony, satire, sarcasm, and cynicism), and the self-defeating scale seems inappropriate to measure humor (or at least the humor style it was intended to measure). Overall, it would be fruitful to focus on comprehensive approaches to humor styles or humor behaviors. Humor does have the potential to benefit both ourselves and other people–we just need to uncover this effect in a psychometrically sound way. This will certainly be a laborious and complex process, but it will likely lead to sustainable and impactful research findings, which in turn enable sound and effective applications in a wide range of areas.

References

Allport, G. W. (1961). *Pattern and growth in personality.* New York, NY: Holt, Rinehart & Winston.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (5th ed.)*. Arlington, VA: American Psychiatric Association.

American Psychological Association. (1954). *Technical recommendations for psychological tests and diagnostic techniques.* Washington, DC: American Psychological Association.

Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*, 1173–1182. doi:10.1037/0022-3514.51.6.1173

Beatty, P. C., & Willis, G. B. (2007). Research synthesis: The practice of cognitive interviewing. *Public Opinion Quarterly, 71*(2), 287–311. doi:10.1093/poq/nfm006

Beermann, U., & Ruch, W. (2011). Can people really "laugh at themselves?" Experimental and correlational evidence. *Emotion, 11*, 492–501. doi:10.1037/a0023444

Behr, D., Bandilla, W., Kaczmirek, L., & Braun, M. (2014). Cognitive probes in Web surveys: On the effect of different text box size and probing exposure on response quality. *Social Science Computer Review, 32*, 524–533. doi:10.1177/0894439313485203

Behr D., Kaczmirek L., Bandilla W., & Braun M. (2012). Asking probing questions in web surveys: Which factors have an impact on the quality of responses? *Social Science Computer Review, 30*, 487–498. doi:10.1177/0894439311435305

Bosson, J. K., Swann Jr., W. B., & Pennebaker, J. W. (2000). Stalking the perfect measure of implicit self-esteem: The blind men and the elephant revisited? *Journal of Personality and Social Psychology* 79. 631–643. doi:10.1037/0022-3514.79.4.631

Buss, A. H., & Finn, S. E. (1987). Classification of personality traits. *Journal of Personality and Social Psychology, 52*, 432–444. doi:10.1037/0022-3514.52.2.432

Caird, S., & Martin, R. A. (2014). Relationship-focused humor styles and relationship satisfaction in dating couples: A repeated-measures design. *Humor: International Journal of Humor Research, 27*, 227–247. doi:10.1515/humor-2014-0015

Caldwell, T. L., Cervone, D., & Rubin, L. H. (2008). Explaining intra-individual variability in social behavior through idiographic assessment: The case of humor. *Journal of Research in Personality, 42*, 1229–1242. doi:10.1016/j.jrp.2008.03.008

Çalisandemir, F., & Tagay, Ö. (2015). Multidimensional perfectionism and humor styles the predictors of life satisfaction. *Procedia-Social and Behavioral Sciences, 174*, 939–945. doi:10.1016/j.sbspro.2015.01.715

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56*, 81–105. doi:10.1037/h0046016

Campos, B., Shiota, M. N., Keltner, D., Gonzaga, G. C., & Goetz, J. L. (2013). What is shared, what is different? Core relational themes and expressive displays of eight positive emotions. *Cognition & Emotion, 27*, 37–52. doi:10.1080/02699931.2012.683852

Cann, A., & Collette, C. (2014). Sense of humor, stable affect, and psychological well-being. *Europe's Journal of Psychology, 10*, 464–479. doi:10.5964/ejop.v10i3.746

Cann, A., Zapata, C. L., & Davis, H. B. (2009). Positive and negative styles of humor in communication: Evidence for the importance of considering both styles. *Communication Quarterly, 57*, 452–468. doi:10.1080/01463370903313398

Cann, A., Zapata, C. L., & Davis, H. B. (2011). Humor styles and relationship satisfaction in dating couples: Perceived versus self-reported humor styles as predictors of satisfaction. *Humor: International Journal of Humor Research*, 24, 1–20. doi:10.1515/HUMR.2011.001

Carretero Dios, H., Eid, M., & Ruch, W. (2011). Analyzing multitrait-mulitmethod data with multilevel confirmatory factor analysis: An application to the validation of the State-Trait Cheerfulness Inventory. *Journal of Research in Personality, 45*, 153–164. doi:10.1016/j.jrp.2010.12.007

Carretero Dios, H., Pérez, C., & Buela Casal, G. (2009). Content validity and metric properties of a pool of items developed to assess humor appreciation. *The Spanish Journal of Psychology, 12*, 773–787. doi:10.1017/S1138741600002146

Carretero Dios, H., Eid, M., & Ruch, W. (2011). Analyzing multitrait-mulitmethod data with multilevel confirmatory factor analysis: An application to the validation of the State-Trait Cheerfulness Inventory. *Journal of Research in Personality, 45*, 153–164. doi:10.1016/j.jrp.2010.12.007

Cattell, R. B., & Luborsky, L. B. (1947). Personality factors in response to humor. *The Journal of Abnormal and Social Psychology, 42*, 402–421. doi:10.1037/h0062651

Cattell, R. B., & Tollefson, D. L. (1966). *IPAT humor test of personality.* Champaign, IL: Institute for Personality and Ability Testing.

Clark, L. A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment, 7*, 309–319. doi:10.1037/1040-3590.7.3.309

Connolly, J. J., Kavanagh, E. J., & Viswesvaran, C. (2007). The convergent validity between self and observer ratings of personality: A meta-analytic review. *International Journal of Selection and Assessment, 15*, 110–117. doi:10.1111/j.1468-2389.2007.00371.x

Coopersmith, S. (1981). *Self-esteem inventories.* Palo Alto, CA: Consulting Psychologist Press.

Craik, K. H., Lampert, M. D., & Nelson, A. J. (1996). Sense of humor and styles of everyday humorous conduct. *Humor: International Journal of Humor Research, 9*, 273–302. doi:10.1515/humr.1996.9.3-4.273

Craik, K. H., & Ware, A. P. (2007). Humor and personality in everyday life. In W. Ruch (Ed.), *The sense of humor: Explorations of a personality characteristic* (pp. 63–94). Berlin, Germany: Mouton de Gruyter.

Crawford, S. A. & Caltabiano, N. J. (2011). Promoting emotional well-being through the use of humour. *The Journal of Positive Psychology, 6*, 237–252. doi:10.1080/17439760.2011.577087

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin, 52*, 281–302. doi:10.1037/h0040957

Delgado Rico, E., Carretero Dios, H., & Ruch, W. (2012). Content validity evidences in test development: An applied perspective. *International Journal of Clinical and Health Psychology, 12*, 449–459. Retrieved from http://www.aepc.es/ijchp/articulos.php?coid=English&id=421

DeYoung, C. G., Carey, B. E., Krueger, R. F., & Ross, S. R. (2016). Ten aspects of the Big Five in the Personality Inventory for DSM–5. *Personality Disorders: Theory, Research, and Treatment, 7*, 113–123. doi:10.1037/per0000170.

Diener, E., & Chan, M. Y. (2011). Happy people live longer: Subjective well-being contributes to health and longevity. *Applied Psychology: Health and Well-Being, 3*, 1–3. doi:10.1111/j.1758-0854.2010.01045.x

Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The satisfaction with life scale. *Journal of Personality Assessment, 49*, 71–75. doi:10.1207/s15327752jpa4901_13

Diener, E., Suh, E. M., Lucas, R. E., & Smith, H. L. (1999). Subjective well-being: Three decades of progress. *Psychological Bulletin, 125*, 276–302. doi:10.1037/0033-2909.125.2.276

Diener, E., Wirtz, D., Tov, W., Kim-Prieto, C., Choi, D. W., Oishi, S., & Biswas-Diener, R. (2010). New well-being measures: Short scales to assess flourishing and positive and negative feelings. *Social Indicators Research, 97*, 143–156. doi:10.1007/sll205-009-9493-y

Donnellan, M. B., Ackerman, R. A., & Brecheen, C. (2016). Extending structural analyses of the Rosenberg self-esteem scale to consider criterion-related validity: Can composite self-esteem scores be good enough? *Journal of Personality Assessment, 98*, 169–177. doi:10.1080/00223891.2015.1058268

Dyck, K. T., & Holtzman, S. (2013). Understanding humor styles and well-being: The importance of social relationships and gender. *Personality and Individual Differences, 55*, 53–58. doi:10.1016/j.paid.2013.01.023

Edwards, K. R., & Martin, R. A. (2010). Humor creation ability and mental health: Are funny people more psychologically healthy? *Europe's Journal of Psychology, 6*, 196–212. doi:10.5964/ejop.v10i3.759

Edwards, K. R., & Martin, R. A. (2014). The conceptualization, measurement, and role of humor as a character strength in positive psychology. *Europe's Journal of Psychology, 10*, 505–519. doi:10.5964/ejop.v10i3.759

Eid, M. (2000). A multitrait-multimethod model with minimal assumptions. *Psychometrika 65*, 241–261. doi:10.1007/BF02294377

Eid, M., Lischetzke, T., Nussbeck, F. W., & Trierweiler, L. I. (2003). Separating trait specific effects from trait-specific method effects in multitrait-multimethod models: A multiple-indicator CT-C(M-1) model. *Psychological Methods, 8*, 38–69. doi:10.1037/1082-989X.8.1.38

Eid, M., Nussbeck, F. W., Geiser, C., Cole, D. A., Gollwitzer, M., & Lischetzke, T. (2008). Structural equation modeling of multitrait-multimethod data: Different models for different types of methods. *Psychological Methods, 13*, 230–253. http://dx.doi.org/10.1037/a0013219

Ekman, P. (2003). *Emotions revealed: Recognizing faces and feelings to improve communication and emotional life.* New York, NY: Times Books.

Ekman, P., Friesen, W. V., & Ancoli, S. (1980). Facial signs of emotional experience. *Journal of Personality and Social Psychology, 39*, 1125–1134. doi:10.1037/h0077722

Ekman, P., Friesen, W. V., & Hager, J. C. (2002). *Facial Action Coding System: The manual.* Salt Lake City, UT: Research Nexus.

Ekman, P., Irwin, W., & Rosenberg, E. L. (1994). *EMFACS-8: Coders instructions.* Unpublished manuscript, University of California, San Diego.

Ekman, P., & Rosenberg, E. L. (2005). *What the face reveals: Basic and applied studies of spontaneous expression using the facial action coding system (FACS).* (2nd ed.) New York, NY: Oxford University Press

Eysenck, H. J. (1942). The appreciation of humour: An experimental and theoretical study. *British Journal of Psychology, 32*, 295–309. doi:10.1111/j.2044-8295.1942.tb01027.x

Eysenck, H. J. (1943). An experimental analysis of five tests of "appreciation of humor". *Educational and Psychological Measurement, 3*, 190–214. doi:10.1177/001316444300300119

Eysenck, H. J. (1944). National differences in "sense of humor": Three experimental and statistical studies. *Journal of Personality, 13*, 37–54. doi:10.1111/j.1467-6494.1944.tb01971.x

Fiedler, K., Schott, M., & Meiser, T. (2011). What mediation analysis can (not) do. *Journal of Experimental Social Psychology, 47*, 1231–1236. doi:10.1016/j.jesp.2011.05.007

Findlay, B. M., & Jones, C. (2005). I think I've got a sense of humour, but do others think so? *Australian Journal of Psychology, 57*, S1, 204. doi:10.1080/00049530600940010

Ford, T. E., Lappi, S. K., & Holden, C. J. (2016). Personality, humor styles and happiness: Happy people have positive humor styles. *Europe's Journal of Psychology, 12*, 320–337, doi:10.5964/ejop.v12i3.1160

Ford, T. E., Lappi, S. K., O'Connor, E. C., & Banos, N. C. (2017). Manipulating humor styles: Engaging in self-enhancing humor reduces state anxiety. *Humor: International Journal of Humor Research*, 30, 169–192. doi:10.1515/humor-2016-0113

Fox, C. L., Dean, S., & Lyford, K. (2013). Development of a humor styles questionnaire for children. *Humor: International Journal of Humor Research, 26*, 295–319. doi:10.1515/humor-2013-0018

Fredrickson, B. L. (2001). The role of positive emotions in positive psychology: The broaden-and-build theory of positive emotions. *American Psychologist, 56*, 218–226. doi:10.1037/0003-066X.56.3.218

Fredrickson, B. L., Tugade, M., Waugh, C., & Larkin, G. (2003). What good are positive emotions in crisis? A prospective study of resilience and emotions following the terrorist attacks on

the United States on September 11th, 2001. *Journal of Personality and Social Psychology, 84*, 365–376. doi:10.1037/0022-3514.84.2.365

Funder, D. C., Furr, R. M., & Colvin, C. R. (2000). The Riverside Behavioral Q-sort: A tool for the description of social behavior. *Journal of Personality, 68*, 451–489. doi:10.1111/1467-6494.00103

Freud, S. (1928). Humour. *International Journal of Psychoanalysis, 9*, 1–6. Retrieved from http://www.pep-web.org/document.php?id=ijp.009.0001a

Frewen, P. A., Brinker, J., Martin, R. A., & Dozois, D. J. A. (2008). Humor styles and personality-vulnerability to depression. *Humor: International Journal of Humor Research,* 21, 179–195. doi:10.1515/HUMOR.2008.009

Furr, R. M. (2009). Personality psychology as a truly behavioural science. *European Journal of Personality, 23*, 369–401. doi:10.1002/per.724

Galloway, G. (2010). Individual differences in personal humor styles: Identification of prominent patterns and their associates. *Personality and Individual Differences, 48*, 563–567. doi:10.1016/j.paid.2009.12.00

Gignac, G. E., & Szodorai, E. T. (2016). Effect size guidelines for individual differences researchers. *Personality and Individual Differences, 102*, 74–78. doi:10.1016/j.paid.2016.06.069

Goldberg, L. R. (2006). Doing it all bass-ackwards: The development of hierarchical factor structures from the top down. *Journal of Research in Personality, 40*, 347–358. doi:10.1016/j.jrp.2006.01.001

Gore, W. L., & Widiger, T. A. (2013). The DSM-5 dimensional trait model and five-factor models of general personality. *Journal of Abnormal Psychology, 122*, 816–821. doi:10.1037/a0032822

Guenter, H., Schreurs, B., Van Emmerik, H., Gijsbers, W., & Van Iterson, A. (2013). How adaptive and maladaptive humor influence well-being at work: A diary study. *Humor: International Journal of Humor Research, 26*, 573–594. doi:10.1515/humor-2013-0032

Haigler, E. D., & Widiger, T. A. (2001). Experimental manipulation of NEO-PI-R items. *Journal of Personality Assessment, 77*, 339–358. doi:10.1207/S15327752JPA7702_14

Heintz, S. (2017). Putting a spotlight on daily humor behaviors: Dimensionality and relationships with personality, subjective well-being, and humor styles. *Personality and Individual Differences, 104*, 407–412. doi:10.1016/j.paid.2016.08.042

Heintz, S., & Ruch, W. (2015). An examination of the convergence between the conceptualization and the measurement of humor styles: A study of the construct validity of the Humor Styles Questionnaire. *Humor: International Journal of Humor Research, 28*, 611–633. doi:10.1515/humor-2015-0095

Heintz, S., & Ruch, W. (2016). Reply to Martin (2015): Why our conclusions hold. *Humor: International Journal of Humor Research, 29*, 125–129. doi:10.1515/humor-2015-0132

Hiranandani, N. A., & Yue, X. D. (2014). Humour styles, gelotophobia and self‐esteem among Chinese and Indian university students. *Asian Journal of Social Psychology, 17*, 319–324. doi:10.1111/ajsp.12066

Hirsh, J. B., & Peterson, J. B. (2009). Personality and language use in self-narratives. *Journal of Research in Personality, 43*, 524–527. doi:10.1016/j.jrp.2009.01.006

Hudson, W. W. (1982). *The clinical measurement package: A field manual.* Chicago, IL: Dorsey Press.

Huber, M., Knottnerus, J. A., Green, L., van der Horst, H., Jadad, A. R., Kromhout, D., … Schnabel, P. (2011). How should we define health? *BMJ: British Medical Journal, 343:* d4163. doi:10.1136/bmj.d4163

Huta, V., & Waterman, A. S. (2014). Eudaimonia and its distinction from hedonia: Developing a classification and terminology for understanding conceptual and operational definitions. *Journal of Happiness Studies, 15*, 1425–1456. doi:10.1007/s10902-013-9485-0

Izard, C. E. (1971). *The face of emotion.* New York, NY: Appleton-Century-Crofts.

Jackson, D. N. (1970). A sequential system for personality scale development. In C.D. Spielberger (Ed.), *Current topics in clinical and community psychology* (Vol.2, pp. 61–96). New York, NY: Academic Press.

James, L., & Fox, C. (2016). The development of a humor styles questionnaire for younger children. *Humor: International Journal of Humor Research, 29*, 555–582. doi:10.1515/humor-2016-0042

Janes, L., & Olson, J. (2015). Humor as an abrasive or a lubricant in social situations: Martineau revisited. *Humor: International Journal of Humor Research, 28*, 271–288. doi:10.1515/humor-2015-0021

Jovanovic, V. (2011). Do humor styles matter in the relationship between personality and subjective well-being? *Scandinavian Journal of Psychology, 52*, 502–507. doi:10.1111/j.1467-9450.2011.00898.x

Kahneman, D., Diener, E., & Schwarz, N. (Eds.). (1999). *Well-being: Foundations of hedonic psychology.* New York, NY: Russell Sage Foundation.

Kazarian, S., Ruch, W., & Proyer, R. T. (2009). Gelotophobia in the Lebanon: The Arabic version of a questionnaire for the subjective assessment of the fear of being laughed at. *Arab Journal of Psychiatry, 20*, 42–56. Retrieved from http://arabpsynet.com/Journals/ajp/ajp20.1-P42.pdf

Kern, M. L., Waters, L. E., Adler, A., & White, M. A. (2015). A multidimensional approach to measuring well-being in students: Application of the PERMA framework. *The Journal of Positive Psychology, 10*, 262–271. doi:10.1080/17439760.2014.936962

Keyes, C. L. M. (1998). Social well-being. *Social Psychology Quarterly, 61*, 121–140. doi:10.2307/2787065

Kirsh, G. A., & Kuiper, N. A. (2003). Positive and negative aspects of sense of humor: Associations with the constructs of individualism and relatedness. *Humor: International Journal of Humor Research, 16*, 33–62. doi:10.1515/humr.2003.004

Kline, R. B. (2015). The mediation myth. *Basic and Applied Social Psychology, 37*, 202–213. doi:10.1080/01973533.2015.1049349

Koch, T., Schultze, M., Burrus, J., Roberts, R. D., & Eid, M. (2015). A multilevel CFA-MTMM model for nested structurally different methods. *Journal of Educational and Behavioral Statistics, 40*, 477–510. doi:10.3102/1076998615606109

Koch, T., Schultze, M., Eid, M., & Geiser, C. (2014). A longitudinal multilevel CFA-MTMM model for interchangeable and structurally different methods. *Frontiers in Psychology, 5*: 311. doi:10.3389/fpsyg.2014.00311

Krohne, H., Egloff, B., Kohlmann, C.-W., & Tausch, A. (1996). Investigations with a German version of the Positive and Negative Affect Schedule (PANAS). *Diagnostica, 42*, 139–156. Retrieved from http://www.hogrefe.de/produkte/zeitschriften/diagnostica

Krueger, R. F., Derringer, J., Markon, K. E., Watson, D., & Skodol, A. E. (2012). Initial construction of a maladaptive personality trait model and inventory for DSM-5. *Psychological Medicine, 42*, 1879-1890. doi:10.1017/S0033291711002674

Krueger, R. F., Derringer, J., Markon, K. E., Watson, D., & Skodol, A. E. (2013). *The personality inventory for DSM-5—brief form (PID-5-BF)—adult.* Arlington, VA: American Psychiatric Association.

Krueger, R. F., Eaton, N. R., Clark, L. A., Watson, D., Markon, K. E., Derringer, J., … Livesley, W. J. (2011). Deriving an empirical structure of personality pathology for DSM-5. *Journal of Personality Disorders, 25*, 170–191. doi:10.1521/pedi.2011.25.2.170

Kuiper, N. (2014, Ed.). Special Issue on "Humor, Well-Being and Health", *Europe`s Journal of Psychology, 10*(3). Retrieved from http://ejop.psychopen.eu/issue/view/56

Kuiper, N. A., Grimshaw, M., Leite, C., & Kirsh, G. (2004). Humor is not always the best medicine: Specific components of sense of humor and psychological well-being. *Humor: International Journal of Humor Research, 17,* 135–168. doi:10.1515/humr.2004.002

Kuiper, N. A., Kirsh, G. A., & Leite, C. (2010). Reactions to humorous comments and implicit theories of humor styles. *Europe's Journal of Psychology, 6*, 236–266. doi:10.5964/ejop.v6i3.215

Kuiper, N., Kirsh, G., & Maiolino, N. (2016). Identity and intimacy development, humor styles, and psychological well-being, *Identity, 16*, 115–125. doi:10.1080/15283488.2016.1159964

Latcheva, R. (2011). Cognitive interviewing and factor-analytic techniques: A mixed method approach to validity of survey items measuring national identity. *Quality & Quantity, 45*, 1175–1199. doi:10.1007/s11135-009-9285-0

Lazarus, R. S. (1991). Progress on a cognitive-motivational-relational theory of emotion. *American Psychologist, 46*, 819–834. http://dx.doi.org/10.1037/0003-066X.46.8.819

Leary, M. R., & Baumeister, R. F. (2000). The nature and function of self-esteem: Sociometer theory. *Advances in Experimental Social Psychology, 32*, 1–62. doi:10.1016/S0065-2601(00)80003-9

Leary, M. R., & MacDonald, G. (2003). Individual differences in self-esteem: A review and theoretical integration. In M. R. Leary, & J. P. Tangney (Eds.), *Handbook of self and identity* (pp. 401–418). New York, NY: Guilford Press.

Leist, A. K., & Müller, D. (2013). Humor types show different patterns of self-regulation, self-esteem, and well-being. *Journal of Happiness Studies, 14*, 551–569. doi:10.1007/s10902-012-9342-6

Lewis, M., Haviland-Jones, J. M., & Barrett, L. F. (Eds.). (2010). *Handbook of emotions.* New York, NY: Guilford Press.

Lodi-Smith, J., & Roberts, B. W. (2007). Social investment and personality: A meta-analysis of the relationship of personality traits to investment in work, family, religion, and volunteerism. *Personality and Social Psychology Review, 11*, 68–86. doi:10.1177/1088868306294590

Loevinger, J. (1957). Objective tests as instruments of psychological theory. *Psychological Reports, 3*, 635–694. doi:10.2466/PR0.3.7.635-694

Luborsky, L. B., & Cattell, R. B. (1947). The validation of personality factors in humor. *Journal of Personality, 15*, 283–291. doi:10.1111/j.1467-6494.1947.tb01068.x

Lucas, R. E., & Dyrenforth, P. S. (2006). Does the existence of social relationships matter for subjective well-being? In K. D. Vohs, & E. J. Finkel (Eds.), Self and relationships: Connecting intrapersonal and interpersonal processes (pp. 254–273). New York, NY: Guilford Press.

Maiolino, N. B., & Kuiper, N. A. (2014). Integrating humor and positive psychology approaches to psychological well-being. *Europe's Journal of Psychology, 10*, 557–570. doi:10.5964/ejop.v10i3.753

Martin, R. A. (2015). On the challenges of measuring humor styles: Response to Heintz and Ruch. *Humor: International Journal of Humor Research, 28,* 635–639. doi:10.1515/humor-2015-0096

Martin, R. A., Puhlik-Doris, P., Larsen, G., Gray, J., & Weir, K. (2003). Individual differences in uses of humor and their relation to psychological well-being: Development of the Humor Styles Questionnaire. *Journal of Research in Personality, 37*, 48–75. doi:10.1016/S0092-6566(02)00534-2

Matsumoto, D., Keltner, D., Shiota, M. N., O'Sullivan, M., & Frank, M. (2010). Facial expression of emotion. In M. Lewis, J. M. Haviland-Jones, & L. F. Barrett (Eds.), *Handbook of emotions* (pp. 211–234, 3rd ed.). New York, NY: Guilford Press.

McGhee, P. E. (1999). *Health, healing and the amuse system: Humor as survival training* (3rd ed.). Dubuque, IA: Kendall/Hunt.

McGhee, P. E. (2010). *Humor as survival training for a stressed-out world: The 7 humor habits program*. Bloomington, IN: AuthorHouse.

McGhee, P. E., & Goldstein, J. H. (Eds.). (1983). *Handbook of humor research, Vol. 1: Basic Issues*. New York, NY: Springer.

McGraw, K. O., & Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods 1*, 30–46. doi:10.1037/1082-989X.1.1.30

Meitinger, K., & Behr, D. (2016). Comparing cognitive interviewing and online probing: Do they find similar results? *Field Methods, 28*, 363–380. doi:10.1177/1525822X15625866

Mendiburo-Seguel, A., Páez, D., & Martínez-Sánchez, F. (2015). Humor styles and personality: A meta-analysis of the relation between humor styles and the Big Five personality traits. *Scandinavian Journal of Psychology, 56*, 335–340. doi:10.1111/sjop.12209

Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist, 50*, 741–749. doi:10.1037/0003-066X.50.9.741.

Meyer, N. A., Helle, A. C., Tucker, R. P., Lengel, G. J., DeShong, H. L., Wingate, L. R., & Mullins-Sweatt, S. N. (2017). Humor styles moderate borderline personality traits and suicide ideation. *Psychiatry Research, 249*, 337–342. doi:10.1016/j.psychres.2017.01.038

Müller, L., & Ruch, W. (2011). Humor and strengths of character. *The Journal of Positive Psychology, 6*, 368–376. doi:10.1080/17439760.2011.592508

Muthén, L. K., & Muthén, B. O. (1998–2008). *Mplus 5.1. Statistical analysis with latent variables. User's guide*. Los Angeles, CA: Muthén and Muthén.

Ozer, D. J. (1993). The Q-sort method and the study of personality development. In D. J. Ozer, , D. C. Funder, R. D. Parke, C. Tomlinson-Keasey, & K. Widaman (Eds). *Studying lives through time: Personality and development* (pp. 147–168). Washington, DC: American Psychological Association. doi:10.1037/10127-023

Ozyesil, Z. (2012). The prediction level of self-esteem on humor style and positive-negative affect. *Psychology, 3*, 638–641. doi:10.4236/psych.2012.38098.

Paulhus, D. L. (2014). Toward a taxonomy of dark personalities. *Current Directions in Psychological Science, 23*, 421–426. doi:10.1177/0963721414547737

Paulhus, D. L., & Williams, K. M. (2002). The dark triad of personality: Narcissism, Machiavellianism, and psychopathy. *Journal of Research in Personality, 36*, 556–563. doi:10.1016/S0092-6566(02)00505-6

Pavot, W., & Diener, E. (1993). Review of the satisfaction with life scale. *Psychological Assessment, 5*, 164–172. doi:10.1037/1040-3590.5.2.164

Pavot, W., & Diener, E. (2013). Happiness experienced: The science of subjective well-being. In S. David, I. Boniwell, & A. Conley Ayers (Eds.), *The Oxford handbook of happiness* (pp. 134–151). New York, NY: Oxford University Press.

Pennebaker, J. W., Francis, M. E., & Booth, R. J. (2001). *Linguistic Inquiry and Word Count – LIWC2001*. Mahwah, NJ: Erlbaum.

Peterson, C., Park, N., & Seligman, M. E. (2005). Orientations to happiness and life satisfaction: The full life versus the empty life. *Journal of Happiness Studies, 6*, 25–41. doi:10.1007/s10902-004-1278-z

Peterson, C., & Seligman, M. E. (2004). *Character strengths and virtues: A handbook and classification.* New York, NY: Oxford University Press.

Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology, 88*, 879–903. doi:10.1037/0021-9010.88.5.879

Presser, S., Rothgeb, J. M., Couper, M. P., Lessler, J. T., Martin, E., Martin, J., & Singer, E. (Eds.) (2004). *Methods for testing and evaluating survey questionnaires.* Hoboken, NJ: John Wiley & Sons.

Rnic, K., Dozois, D. J. A., & Martin, R. A. (2016). Cognitive distortions, humor styles, and depression. *Europe's Journal of Psychology, 12*. 348–362. doi:10.5964/ejop.v12i3.1118

Rosenberg, M. (1965). *Society and the adolescent self-image.* Princeton, NJ: Princeton University Press.

Rosenberg, M., Schooler, C., & Schoenbach, C. (1989). Self-esteem and adolescent problems: Modeling reciprocal effects. *American Sociological Review, 54*, 1004–1018. doi:10.2307/2095720

Rosenberg, M., Schooler, C., Schoenbach, C., & Rosenberg, F. (1995). Global self-esteem and specific self-esteem: Different concepts, different outcomes. *American Sociological Review, 60*, 141–156. doi:10.2307/2096350

Ruch, W. (1998). Foreword and overview: Sense of humor: A new look at an old concept. In W. Ruch (Ed.), *The sense of humor: Explorations of a personality characteristic* (pp. 3–14). Berlin, Germany: Mouton de Gruyter.

Ruch, W. (2012). Towards a new structural model of the sense of humor: Preliminary findings. In V. Raskin & J. M. Taylor (Chairs), *Artificial Intelligence of Humor: Papers from the 2012 AAAI Fall Symposium* (pp. 68–75). Retrieved from http://www.ilhaire.eu/pdf/Ruch_2012.pdf

Ruch, W., & Heintz, S. (2013). Humour styles, personality, and psychological well-being: What's humour got to do with it? *European Journal of Humour Research, 1*(4), 1–24. doi:10.7592/EJHR2013.1.4.ruch

Ruch, W., & Heintz, S. (2016a). The virtue gap in humor: Exploring benevolent and corrective humor. *Translational Issues in Psychological Science, 2*, 35–45. doi:10.1037/tps0000063

Ruch, W., & Heintz, S. (2016b). The German version of the Humor Styles Questionnaire: Psychometric properties and overlap with other styles of humor. *Europe's Journal of Psychology, 12*, 434–455. doi:10.5964/ejop.v12i3.1116

Ruch, W., & Heintz, S. (2017). Experimentally manipulating items informs on the (limited) construct and criterion validity of the Humor Styles Questionnaire. *Frontiers in Psychology: Personality and Social Psychology, 8*: 616. doi:10.3389/fpsyg.2017.00616

Ruch, W., Heintz, S., Platt, T., Proyer, R. T., & Wagner, L. (2018). Broadening humor: Comic styles differentially tap into temperament, character, and ability. *Frontiers in Psychology: Personality and Social Psychology, 9*: 6. doi:10.3389/fpsyg.2018.00006

Ruch, W., & Hofmann, J. (2017). Fostering humour. In C. Proctor (Ed.), *Positive psychology interventions in practice* (pp. 65–80). New York, NY: Springer International Publishing. doi:10.1007/978-3-319-51787-2_5

Ruch, W., Hofmann, J., Rusch, S., & Stolz, H. (2018). Training the sense of humor with the 7 Humor Habits Program and satisfaction with life. *Humor: International Journal of Humor Research, 31*, 287–309. doi:10.1515/humor-2017-0099

Ruch, W., Köhler, G., & Van Thriel, C. (1996). Assessing the „humorous temperament “: Construction of the facet and standard trait forms of the State-Trait-Cheerfulness-Inventory—STCI. *Humor: International Journal of Humor Research, 9*, 303–340. doi:10.1515/humr.1996.9.3-4.303

Ruch, W., & McGhee, P. E. (2014). Humor intervention programs. In Acacia C. Parks & Stephen M. Schueller (Eds.), *The Wiley Blackwell Handbook of Positive Psychological Interventions* (pp. 179–193). Chichester, UK: John Wiley & Sons.

Ruch, W., Proyer, R. T., Esser, C., & Mitrache, O. (2011). Cheerfulness and everyday humorous conduct. In Romanian Academy (Ed.), *Yearbook of the 'George Baritiu' Institute of History in Cluj-Napoca – Series Humanistica* (vol. 9, pp. 67–87). Bucharest, Romania: Romanian Academy Publishing House.

Rusk, R. D., & Waters, L. (2015). A psycho-social system approach to well-being: Empirically deriving the five domains of positive functioning. *The Journal of Positive Psychology, 10*, 141–152. doi:10.1080/17439760.2014.920409

Ryan, R. M., & Deci, E. L. (2001). On happiness and human potentials: A review of research on hedonic and eudaimonic well-being. *Annual Review of Psychology, 52*, 141–166. doi:10.1146/annurev.psych.52.1.141

Ryff, C. D. (1989). Happiness is everything, or is it? Explorations on the meaning of psychological well-being. *Journal of Personality and Social Psychology, 57*, 1069–1081. doi:10.1037/0022-3514.57.6.1069

Ryff, C. D., & Singer, B. H. (2013). Know thyself and become what you are: A eudaimonic approach to psychological well-being. In A. Delle Fave (Ed.), *The exploration of happiness: Present and future perspectives* (pp. 97–116). Dordrecht, Netherlands: Springer.

Saroglou, V., & Scariot, C. (2002). Humor Styles Questionnaire: Personality and educational correlates in Belgian high school and college students. *European Journal of Personality, 16*, 43–54. doi:10.1002/per.430

Schallberger, U., & Venetz, M. (1999). *Kurzversionen des MRS-Inventars von Ostendorf (1990) zur Erfassung der fünf "grossen" Persönlichkeitsfaktoren [Brief versions of Ostendorf's (1990) MRS inventory for the assessment of the Big Five personality factors]*. Zurich, Switzerland: Department of Applied Psychology, University of Zurich.

Scheel, T., Gerdenitsch, C., & Korunka, C. (2016). Humor at work: Validation of the short work-related Humor Styles Questionnaire (swHSQ). *Humor: International Journal of Humor Research, 29*, 439–465. doi:10.1515/humor-2015-0118

Schmidt-Hidding, W. (1963). *Europäische Schlüsselwörter. Band 1: Humor und Witz [European key words. Vol I: Humor and wit]*. Munich, Germany: Huber.

Schneider, L., & Schimmack, U. (2009). Self-informant agreement in well-being ratings: A meta-analysis. *Social Indicators Research, 94*, 363–376. doi:10.1007/s11205-009-9440-y

Schönbrodt, F. D., & Perugini, M. (2013). At what sample size do correlations stabilize? *Journal of Research in Personality, 47*, 609–612. doi:10.1016/j.jrp.2013.05.009

Seligman, M. E. (2012). *Flourish: A visionary new understanding of happiness and well-being.* New York, NY: Free Press.

Shiota, M. N., Campos, B., Oveis, C., Hertenstein, M. J., Simon-Thomas, E., & Keltner, D. (2017). Beyond happiness: Building a science of discrete positive emotions. *American Psychologist, 72*, 617–643. doi:10.1037/a0040456

Simms, L. J. (2008). Classical and modern methods of psychological scale construction. *Social and Personality Psychology Compass, 2*, 414–433. doi:10.1111/j.1751-9004.2007.00044.x

Sirigatti, S., Penzo, I., Giannetti, E., & Stefanile, C. (2014). The Humor Styles Questionnaire in Italy: Psychometric properties and relationships with psychological well-being. *Europe's Journal of Psychology, 10*, 429–450. doi:10.5964/ejop.v10i3.682

Small, M. L. (2011). How to conduct a mixed methods study: Recent trends in a rapidly growing literature. *Annual Review of Sociology, 37*, 57–86. doi:10.1146/annurev.soc.012809.102657

Soto, C. J. (2015). Is happiness good for your personality? Concurrent and prospective relations of the big five with subjective well‐being. *Journal of Personality, 83*, 45–55. doi:10.1111/jopy.12081

Sowislo, J. F., & Orth, U. (2013). Does low self-esteem predict depression and anxiety? A meta-analysis of longitudinal studies. *Psychological Bulletin, 139*, 213–240. doi:10.1037/a0028931

Spearman, C. (1904). The proof and measurement of association between two things. *American Journal of Psychology 15*, 72–101. doi:10.2307/1412159

Steel, P., Schmidt, J., & Shultz, J. (2008). Refining the relationship between personality and subjective well-being. *Psychological Bulletin, 134*, 138–161. doi:10.1037/0033-2909.134.1.138

Stelmack, R. M., & Stalikas, A. (1991). Galen and the humour theory of temperament. *Personality and Individual Differences, 12*, 255–263. doi:10.1016/0191-8869(91)90111-N

Stieger, S., Formann, A. K., & Burger, C. (2011). Humor styles and their relationship to explicit and implicit self-esteem. *Personality and Individual Differences, 50*, 747–750. doi:10.1016/j.paid.2010.11.025

Stokenberga, I. (2008). Humorous personality: Relationship to stress and well-being. *Baltic Journal of Psychology, 9*, 70–80. Retrieved from http://dspace.lu.lv/dspace/bitstream/handle/7/1317/BalticJournPsychol-2008-Vol-9_No-1-2.pdf

Stones, M. J., Worobetz, S., & Brink, P. (2011). Overestimated relationships with subjective well-being. *Canadian Psychology/Psychologie Canadienne, 52*, 93–100. doi:10.1037/a0020694

Su, R., Tay, L., & Diener, E. (2014). The development and validation of the Comprehensive Inventory of Thriving (CIT) and the Brief Inventory of Thriving (BIT). *Applied Psychology: Health and Well‐Being, 6*, 251–279. doi:10.1111/aphw.12027

Tafarodi, R. W., & Milne, A. B. (2002). Decomposing global self‐esteem. *Journal of Personality, 70*, 443–484. doi:10.1111/1467-6494.05017

Tate, C. U. (2015). On the overuse and misuse of mediation analysis: It may be a matter of timing. *Basic and Applied Social Psychology, 37*, 235–246. doi:10.1080/01973533.2015.1062380

Tracy, J. L., Robins, R. W. & Tangney, J. P. (Eds.) (2007). *The self-conscious emotions: Theory and research.* New York, NY: Guilford Press.

Tucker, R. P., Judah, M. R., O'Keefe, V. M., Mills, A. C., Lechner, W. V., Davidson, C. L., … Wingate, L. R. (2013). Humor styles impact the relationship between symptoms of social anxiety and depression. *Personality and Individual Differences, 55*, 823–827. doi:10.1016/j.paid.2013.07.008

Vaughan, J., Zeigler-Hill, V., & Arnau, R. C. (2014). Self-esteem instability and humor styles: Does the stability of self-esteem influence how people use humor? *Journal of Social Psychology, 154*, 299–310. doi:10.1080/00224545.2014.896773

Waterman, A. S. (1990). The relevance of Aristotle's conception of eudaimonia for the psychological study of happiness. *Theoretical & Philosophical Psychology, 10*, 39–44. doi:10.1037/h0091489

Waterman, A. S. (1993). Two conceptions of happiness: Contrasts of personal expressiveness (eudaimonia) and hedonic enjoyment. *Journal of Personality and Social Psychology, 64*, 678–691. doi:10.1037/0022-3514.64.4.678

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology, 54*, 1063–1070. doi:10.1037/0022-3514.54.6.1063

Wellenzohn, S., Proyer, R. T., & Ruch, W. (2016a). Humor-based online positive psychology interventions: A randomized placebo-controlled long-term trial. *Journal of Positive Psychology, 11*, 584–594. doi:10.1080/17439760.2015.1137624

Wellenzohn, S., Proyer, R. T., & Ruch, W. (2016b). How do positive psychology interventions work? A short-term placebo-controlled humor-based study on the role of the time focus. *Personality and Individual Differences, 96*, 1–6. doi:10.1016/j.paid.2016.02.056

Wilt, J., & Revelle, W. (2015). Affect, behaviour, cognition and desire in the Big Five: An analysis of item content and structure. *European Journal of Personality, 29*, 478–497. doi:10.1002/per.2002

Wirth, U. (Ed.) (2017). *Handbuch der Komik [Handbook of the comic]*. Stuttgart, Germany: Springer.

Wolff, H. A., Smith, C. E., & Murray, H. A. (1934). The psychology of humor. *The Journal of Abnormal and Social Psychology, 28*, 341–365. doi:10.1037/h0075400

World Health Organization (2006). *Constitution of the World Health Organization.* Retrieved from www.who.int/governance/eb/who_constitution_en.pdf

Yarkoni, T. (2010). Personality in 100,000 words: A large-scale analysis of personality and word use among bloggers. *Journal of Research in Personality, 44*, 363–373. doi:10.1016/j.jrp.2010.04.001

Yue, X. D., Liu, K. W. Y., Jiang, F., & Hiranandani, N. A. (2014). Humor styles, self-esteem, and subjective happiness. *Psychological Reports, 115*, 517–525. doi:10.2466/07.02.PR0.115c18z6

Zeigler-Hill, V., & Besser, A. (2011). Humor style mediates the association between pathological narcissism and self-esteem. *Personality and Individual Differences, 50*, 1196–1201. doi:10.1016/j.paid.2011.02.006

Zeigler-Hill, V., Besser, A., & Jett, S. E. (2013). Laughing at the looking glass: Does humor style serve as an interpersonal signal? *Evolutionary Psychology, 11*, 201–226. doi:10.1177/147470491301100118

Zeigler‐Hill, V., Besser, A., Myers, E. M., Southard, A. C., & Malkin, M. L. (2013). The status‐signaling property of self‐esteem: The role of self‐reported self‐esteem and perceived self‐esteem in personality judgments. *Journal of Personality, 81*, 209–220. doi:10.1111/j.1467-6494.2012.00790.x

Zeigler-Hill, V., McCabe, G. A., & Vrabel, J. K. (2016). The dark side of humor: DSM-5 pathological personality traits and humor styles. *Europe's Journal of Psychology, 12*, 363–376. doi:10.5964/ejop.v12i3.1109

Zhao, J., Kong, F., & Wang, Y. (2012). Self-esteem and humor style as mediators of the effects of shyness on loneliness among Chinese college students. *Personality and Individual Differences, 52*, 686–690. doi:10.1016/j.paid.2011.12.024

Zhao, J., Wang, Y., & Kong, F. (2014). Exploring the mediation effect of social support and self-esteem on the relationship between humor style and life satisfaction in Chinese college students. *Personality and Individual Differences, 64*, 126–130. doi:10.1016/j.paid.2014.02.026

Ziegler, M. (2014). Stop and state your intentions! Let's not forget the ABC of test construction. *European Journal of Psychological Assessment, 30*, 239–242. doi:10.1027/1015-5759/a000228

Ziegler, M., Booth, T., & Bensch, D. (2013). Getting entangled in the nomological net: Thoughts on validity and conceptual overlap. *European Journal of Psychological Assessment, 29*, 157–161. doi:10.1027/1015-5759/a000173

Zimmermann, J., Krueger, R. F., Markon, K. E., & Leising, D. (2012). *German translation of the personality inventory for DSM-5—brief form (PID-5-BF)—adult.* Unpublished manuscript, University of Kassel, Germany.

Ehrenwort

Hiermit erkläre ich, dass die Dissertation von mir selbst ohne unerlaubte Beihilfe verfasst worden ist.

Ort und Datum Unterschrift

Curriculum Vitae

Sonja Heintz, born on October 10, 1987

Employment history

2018 – current	Postdoctoral researcher in the Swiss National Science Foundation (SNSF) project "The Foundations of a Good Life – Studies on the Advancement of Character Research"
2017 – current	Postdoctoral researcher at the Section of Personality and Assessment, Department of Psychology at the University of Zurich, Zurich, Switzerland
2016 – current	Assistant in the module 07 "Persönlichkeit und Differentielle Psychologie" [Personality and Differential Psychology], Bachelor of Science in Psychology, Swiss Distance Learning University, Brig, Switzerland
2012 – 2017	Research and teaching assistant at the Section of Personality and Assessment, Department of Psychology at the University of Zurich, Zurich, Switzerland
2012 – 2015	Secretary at the Section of Personality and Assessment, Department of Psychology at the University of Zurich, Zurich, Switzerland
2009 – 2011	Student research assistant at the Section of Development of Language, Learning, and Action, Department of Psychology at the Saarland University, Saarbrücken, Germany

Education

2013 – 2017	PhD studies at the Section of Personality and Assessment, Department of Psychology at the University of Zurich, Switzerland (main supervisor: Prof. Willibald Ruch of the University of Zurich, second supervisor: Prof. Dr. Hugo Carretero Dios of the Universidad de Granada)

2007 – 2012 Diploma studies (equivalent to Master of Science) in Psychology at the Saarland University in Saarbrücken, Germany

1998 – 2007 Warndt-Gymnasium in Völklingen, Germany – degree: Abitur (university-entrance diploma)

Publications

Peer-reviewed journal articles [* denotes shared first authorship]

Heintz, S., & Ruch, W. (2018). Can self-defeating humor make you happy? Cognitive interviews reveal the adaptive side of the self-defeating humor style. *Humor: International Journal of Humor Research.* Advance online publication. doi:10.1515/HUMOR-2017-0089

Ruch, W., & Heintz, S. (2018). Psychometric evaluation of the revised Sense of Humor Scale and the construction of a parallel form. *Humor: International Journal of Humor Research: Special Issue "Festschrift for Paul E. McGhee".* Advance online publication. doi:10.1515/HUMOR-2016-0085

Ruch, W., Wagner, L., & Heintz, S. (2018). Humor, the PEN model of personality, and subjective well-being: Support for differential relationships with eight comic styles. *Rivista Italiana di Studi sull'Umorismo, 1*(1), 31–43. Retrieved from www.risu.biz/wp-content/uploads/2018/01/Ruch_et_al.-RISU-11-2018-31-44-1.pdf

Ruch, W., Heintz, S., Platt, T., Wagner, L., & Proyer, R. T. (2018). Broadening humor: Comic styles differentially tap into temperament, character, and ability. *Frontiers in Psychology: Personality and Social Psychology, Research topic "Humor and Laughter, Playfulness and Cheerfulness: Upsides and Downsides to a Life of Lightness", 9*: 6. doi:10.3389/fpsyg.2018.00006

Heintz, S., Kramm, C., & Ruch, W. (2017). A meta-analysis of gender differences in character strengths and age, nation, and measure as moderators. *The Journal of Positive Psychology*. Advance online publication. doi:10.1080/17439760.2017.1414297

Heintz, S. (2017). Do others judge my humor style as I do? Self-other agreement and construct validity of the Humor Styles Questionnaire. *European Journal of Psychological Assessment.* Advance online publication. doi:10.1027/1015-5759/a000440

Ruch, W.*, & Heintz, S.* (2017). Experimentally manipulating items informs on the (limited) construct and criterion validity of the Humor Styles Questionnaire. *Frontiers in Psychology: Personality and Social Psychology, Research topic "Humor and Laughter, Playfulness and Cheerfulness: Upsides and Downsides to a Life of Lightness", 8*:616. doi:10.3389/fpsyg.2017.00616

Heintz, S. (2017). Putting a spotlight on daily humor behaviors: Dimensionality and relationships with personality, subjective well-being, and humor styles. *Personality and Individual Differences, 104*, 407–412. doi:10.1016/j.paid.2016.08.042

Ruch, W.*, & Heintz, S.* (2016). The German version of the Humor Styles Questionnaire: Psychometric properties and overlap with other styles of humor. *Europe's Journal of Psychology: Special Issue "Psychological Investigations of Humor and Laughter. Honoring the Research Contributions of Professor Rod Martin"*, *12*, 434–455. doi:10.5964/ejop.v12i3.1116

Ruch, W., & Heintz, S. (2016). The virtue gap in humor: Exploring benevolent and corrective humor. *Translational Issues in Psychological Science: Special Issue "The Psychology of Humor",* 2, 35–45. doi:10.1037/tps0000063

Heintz, S., & Ruch, W. (2016). Reply to Martin (2015): Why our conclusions hold. *Humor: International Journal of Humor Research*, *29*, 125–129. doi:10.1515/humor-2015-0132

Heintz, S. & Ruch, W. (2015). An examination of the convergence between the conceptualization and the measurement of humor styles: A study of the construct validity of the Humor Styles Questionnaire. *Humor: International Journal of Humor Research, 28*, 611–633. doi:10.1515/humor-2015-0095

Kray, J., Schmitt, H., Heintz, S., & Blaye, A. (2015). Does verbal labeling influence age differences in proactive and reactive cognitive control? *Developmental Psychology, 51*, 378–391. doi:10.1037/a0038795

Ruch, W., Martínez-Martí, M. L., Heintz, S., & Brouwers, S. A. (2014). Short form of the Orientations to Happiness Questionnaire for the German-speaking countries: Development and analysis of the psychometric properties. *Swiss Journal of Psychology, 73*, 225–234. doi:10.1024/1421-0185/a000141

Ruch, W., & Heintz, S. (2013). Humour styles, personality, and psychological well-being: What's humour got to do with it? *European Journal of Humour Research, 1*(4), 1–24. doi:10.7592/EJHR2013.1.4.ruch

Unger, K., Heintz, S., & Kray, J. (2012). Punishment sensitivity modulates the processing of negative feedback but not error-induced learning. *Frontiers in Human Neuroscience, 6:* 186. doi:10.3389/fnhum.2012.00186

Book chapters

Ruch, W., & Heintz, S. (in press). Humor production and creativity: Overview and recommendations. To appear in S. R. Luria, J. Baer, & J. C. Kaufman (Eds.), *Creativity and humor*. San Diego, CA: Academic Press.

Ruch, W., & Heintz, S. (2014). Factor analysis of humor scales. In S. Attardo (Ed.), *Encyclopedia of humor studies (vol. 1, pp. 228–231)*. Thousand Oaks, CA: Sage.

Ruch, W., & Heintz, S. (2014). Test measurements of humor. In S. Attardo (Ed.), *Encyclopedia of humor studies (vol. 2, pp. 759–761)*. Thousand Oaks, CA: Sage.